Chapter 7
Hedonics

Laura O. Taylor

Abstract This chapter covers the current theory and empirical methods in hedonic valuation of environmental and natural resources. A framework is first presented that links hedonic price functions to theoretically correct welfare measures for changes in environmental amenities. The major empirical methods for estimating a hedonic price function are discussed beginning with data construction and basic estimation approaches, and progressing through to techniques for addressing endogenous regressors including spatial econometrics and quasi-experimental methods. The use of the hedonic price function for obtaining measures of welfare change for changes in environmental amenities are also presented. Sorting models and second-stage demand analysis in both a single-market and multiple-market context are described. Applications and examples from housing and labor markets are used throughout to illustrate concepts covered.

Keywords Hedonic method · Implicit prices · First-stage estimation · Second-stage estimation · Quasi-experimental methods · Sorting models · Welfare measures

Heterogeneous or differentiated goods are products whose characteristics vary in such a way that there are distinct product varieties even though the product is sold in one market (e.g., cars, computers, houses). The variation in product variety gives rise to variations in product prices within each market. The hedonic method relies on market transactions for these differentiated goods to determine the implied value or implicit price of characteristics. For instance, by observing the price differential between two product varieties that vary only by one characteristic (e.g., two identical cars, but one has more horsepower than the other), we indirectly observe the monetary trade-offs individuals are willing to make with respect to the changes in this characteristic, and the value of the increase in horsepower is the difference in the prices of the two cars. As such, the hedonic method is an indirect valuation

---

L.O. Taylor (✉)
North Carolina State University, Raleigh, USA
e-mail: laura_taylor@ncsu.edu

P.A. Champ et al. (eds.), *A Primer on Nonmarket Valuation*,
The Economics of Non-Market Goods and Resources 13,
DOI 10.1007/978-94-007-7104-8_7

method where we do not observe the value consumers have for the characteristic directly, but infer it from observable market transactions.

The most common application of hedonic theory in environmental valuation involves housing markets. The choice of housing location and, therefore, neighborhood amenities, is observable. Often, location choice is directly linked to an environmental amenity of interest. For example, housing locations can offer different scenic vistas (Paterson and Boyle 2002; Bin et al. 2008), or they can impose greater perceived risks by placing a household closer to perceived hazards, such as shale gas extraction activities (Gopalakrishnan and Klaiber 2014). As such, the choice of a house and its associated price implies an implicit choice over the environmental amenities (or disamenities) proximate to the house and their implicit prices.

To see how implicit prices for environmental goods are revealed through market transactions, imagine the following scenario in which there are two identical lakes, each with 100 identical homes surrounding them. All homes are lakefront, and all the characteristics of the homes themselves, the land, and the neighborhoods are identical across the properties. At the current equilibrium price of $200,000 per house, all 200 homes are equally preferred. Now, imagine that water clarity at one lake, Lake A for example, is improved. We assume that the improved water clarity is preferred by all households. Now if any home on Lake A were offered at the original equilibrium price of $200,000, consumers would uniformly prefer this house to any house on Lake B. In other words, at the current price, there would be excess demand for the houses located on Lake A, and as such, the price of these houses must rise to bring the market into equilibrium. The price differential that results from the change in water clarity at Lake A is the implicit price consumers are willing to pay for that incremental increase in water clarity. This willingness to pay for water clarity is indirectly revealed to us through the market prices of the homes. For instance, if in the new equilibrium, houses on Lake A sell for $210,000 while houses on Lake B sell for $200,000, the implicit price associated with the increased water clarity is $10,000.

Of course, housing markets aren't so simple: housing choice depends on many characteristics, such as structure of the house, amenities of the land, neighborhood, and location. Yet, the fundamental intuition behind the hedonic method extends easily. By observing the choices consumers make over heterogeneous commodities with varying prices, we can estimate the implicit prices of component characteristics. These implicit prices or hedonic prices, under certain conditions, are equal to Marshallian willingness to pay (WTP) or allow us to recover WTP.

Hedonic analyses have been reported as early as Waugh's (1928) analysis of quality factors influencing asparagus pricing and have been applied to markets as varied as automobiles, computers, VCRs, household appliances, art, and agricultural commodities such as organic produce and wine. This chapter focuses primarily on housing markets and how they can be used to value environmental amenities. The application of the hedonic method to labor markets is also briefly reviewed in Sect. 7.5 (Bockstael and McConnell, 2007, provided a detailed review). Other reviews of the hedonic method can be found in Palmquist (2006), and Phaneuf and Requate (in press).

Hedonic analysis of markets for differentiated goods consists of two related steps often referred to as a first-stage and second-stage analysis. In a first-stage analysis, the hedonic price function is estimated using information about the sales prices of a differentiated product and the characteristics of the product. This analysis allows researchers to recover the implicit prices of characteristics and reveals information on the underlying preferences for these characteristics, which is discussed in Sect. 7.3. First-stage analyses are the most common application of the hedonic method because the needed economic insights often require only marginal price information, and the data are generally readily available. An introduction to the empirical method and data collection is presented in Sects. 7.1 and 7.2.2.

Although econometric estimation of the hedonic price function is conceptually straightforward, housing and other real estate markets have unique challenges. In particular, hedonic housing applications are usually interested in environmental features that vary over space, such as air quality across an urban area. Concerns about unobservable characteristics of housing and their neighborhoods that co-vary over space with environmental features of interest have led to an increased use of spatial econometrics and quasi-experimental methods to address the potential for biased coefficient estimates due to endogenous regressors. Both of these approaches are discussed in Sect. 7.2.3.

Once a first-stage analysis is completed, researchers can combine information on the implicit prices obtained in the first-stage with data on household characteristics to estimate demand functions for the characteristics or utility parameters. This step is referred to as a second-stage analysis and is discussed in Sect. 7.4. While second-stage analysis had traditionally been less common due to data demands, this is changing rapidly as data acquisition costs have decreased. Second-stage analyses are also very important because environmental policy analyses frequently require welfare analyses of large-scale environmental changes, such as air quality changes across an entire metropolitan area, and these situations require that underlying demands or utility parameters be recovered.

Before discussing the implementation of first- and second-stage analyses, the next section reviews the theory that provides a framework for understanding the market process generating a hedonic equilibrium. For the purposes of nonmarket valuation, Rosen's (1974) seminal article provided this theory, and it is important because it developed the utility theoretic framework that establishes the connections between consumers' preferences for characteristics of heterogeneous goods and the equilibrium price function.

7.1 The Hedonic Price Function: Theory

This chapter discusses the hedonic theory using the following notation and assumptions. Let Z represent the differentiated product with characteristics $\underline{z} = z_1, z_2, z_3, \ldots z_n$. The differentiated product is assumed to be sold in a perfectly competitive market, and the interactions of the many producers and consumers together determine an equilibrium price schedule for the differentiated product, $P(\underline{z})$. The equilibrium price for any one particular variety of a differentiated good (e.g., a specific house) is a function of the characteristics of that particular variety. As such, the consumer can determine the price he or she pays for the good by choosing which model to purchase. However, it is important to note that the consumer takes the entire price schedule $P(\underline{z})$ as exogenous.

Consumer utility is defined over two goods: Z, the differentiated good, and x, a composite product representing all other goods (i.e., income left over after purchasing Z). Consumer j, with demographic characteristics α^j, has utility defined as

$$U^j\left(x, z_1, z_2, \ldots, z_n; \alpha^j\right). \tag{7.1}$$

If one assumes that the consumer purchases only one unit of the differentiated product, a reasonable assumption for the housing market, the budget constraint is $y^j = x + P(\underline{z})$. The consumer seeks to maximize utility by choosing the model of the differentiated product, $\underline{z}$, and the amount of x to purchase, subject to his or her budget constraint. The consumer will choose $\underline{z}$ and x such that the following is satisfied for each z_i:

$$\frac{\partial P}{\partial z_i} = \frac{\partial U/\partial z_i}{\partial U/\partial x}, \tag{7.2}$$

which indicates that the marginal rate of substitution between any characteristic, z_i, and the composite numeraire good, x, is equal to the rate at which the consumer can trade z_i for x in the market (i.e., the ratio of the implicit price for z_i and the price of the numeraire, which is, by definition, equal to one).

A second convenient way to describe the optimization process is to define the optimal bid a consumer will make for any specific product variety. The bid function, θ, describes the relationship between the dollar bid Consumer j will make for Z as one or more of its component characteristics are changed while utility and income remain constant. Equation (7.1) can be used to formally define the bid function by recognizing that income less the bid a consumer makes for Z is the amount of money left over to spend on the numeraire, x. Thus, the relationship

$$U^j\left(y_0 - \theta, \underline{z}, \alpha^j\right) \equiv U_0^j, \tag{7.3}$$

indicates how a consumer's optimal bid must vary in response to changes in z if utility and income are held constant, where y is exogenous income, and U_0 is a fixed level of utility. Solving Eq. (7.3) for θ indicates that $\theta^j = \theta(\underline{z}, y_0, U_0^j, \alpha^j)$. Maximizing utility in (7.3) yields the result that the marginal bid a consumer is willing to make for z_i ($\partial\theta/\partial z_i$, which is denoted as θ_i), will equal the marginal rate of substitution between any characteristic, z_i, and x. Relating this result to Eq. (7.2) indicates that the necessary condition for utility maximization is that the marginal bid a consumer places for a characteristic must equal the marginal price of that characteristic ($\partial P(\underline{z})/\partial z_i$, which is denoted as P_i).

For the supply side of the market, one can describe a firm with characteristics δ^k as seeking to maximize profits: $\Pi = H * P(\underline{z}) - C(H, \underline{z}, \delta^k)$, where H is the number of units of Z that the firm produces, and $C(\cdot)$ is a well-behaved cost function. Again, the firm faces the exogenous equilibrium price schedule, $P(\underline{z})$, when determining its

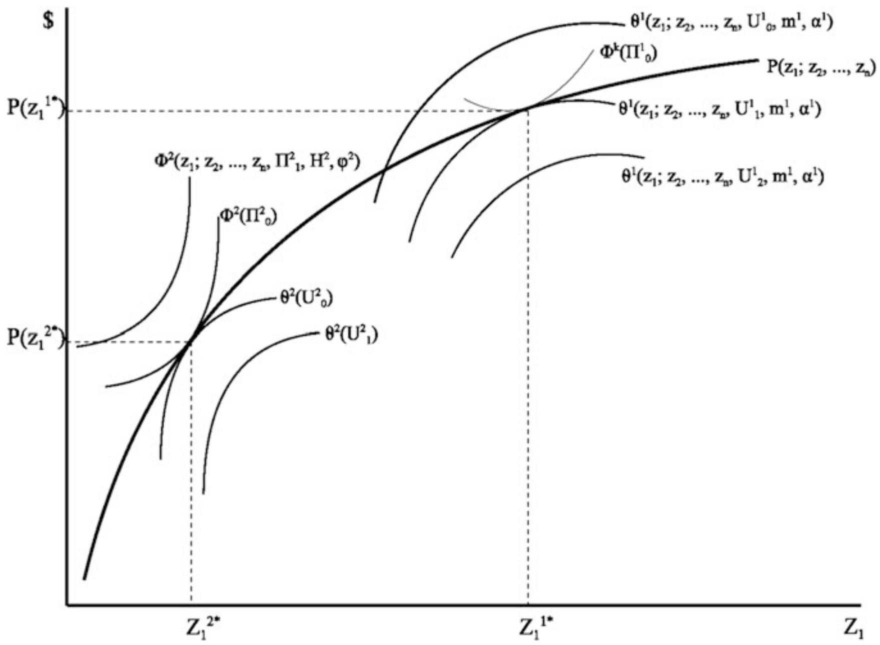

Fig. 7.1 The hedonic price function

choices. Although firms can affect the prices they receive for their products by varying the characteristics of the product, no single firm can affect the price schedule. In this formulation, we assume the firm produces only one model of Z. Thus, the firm chooses what type to produce, Z^k, and then chooses how many of that type to produce.[1] Similar to the consumer's problem, one can describe the behavior of a firm in this differentiated goods market by an offer function, $\varphi^k = \varphi(\underline{z}; H, \Pi_0, \delta^k)$, which describes the amount of money a firm is willing to accept for any particular variety of Z, holding constant the number of units of that variety the firm produces, H, and its level of profit, Π_0. The offer function is defined by $\Pi_0 = H * \varphi^k - C(H, \underline{z}, \delta^k)$, and at the optimum, it will be the case that the marginal price a firm is willing to accept for $\underline{z}_i$, φ_{zi}, will equal the marginal cost of producing that characteristic per unit of the differentiated good, C_{zi}/H.

The bid and offer functions and their properties may be easily described using Fig. 7.1. This figure illustrates the equilibrium price schedule, $P(\underline{z})$ as it varies with changes in z_1, holding the level of all other characteristics constant. $P(\underline{z})$ is drawn such that the total price paid for z_1 increases at a decreasing rate, which one might expect in many applications. For instance, say z_1 represents the number of square

[1]It may also be assumed that firms produce multiple types of the good and that the cost function is separable in each type of product. The maximization problem for the firm choosing an entire product line in this case is equivalent to choosing each product type separately, as described here.

feet of living space in a house. We might expect a smaller price differential between a 5,000- and a 5,300-square-foot house as compared to the price differential between a 1,000- and a 1,300-square-foot house.

Also depicted in Fig. 7.1 are the bid functions for two consumers, θ^1 and θ^2. Along any bid function contour, only the level of z_1 changes; the level of all other characteristics, income, and utility are constant. Bid functions are assumed to be concave in $\underline{z}$ (i.e., optimal bids increase at a decreasing rate in $\underline{z}$), and higher levels of utility for a consumer are represented by bid function contours closer to the horizontal axis. Intuitively, a lower bid for the same level of z_1 implies a higher level of utility because more money is left over to spend on x.[2] The optimal choice of z_1 is where the consumer reaches the lowest possible bid function while still being able to participate in the market. For Consumer 1, this occurs at a quantity z_1^{1*} and a total bid price of $P(z_1^{1*})$, which is the point at which the bid function is tangent to the equilibrium price schedule in Fig. 7.1. For Consumer 2, this optimal choice occurs at z_1^{2*}, and the consumer pays a total price of $\mathrm{P}(z_1^{2*})$. At each consumer's optimal choice of z_1, the marginal bid ($\partial\theta/\partial z_1$) equals the marginal price ($\partial P(\underline{z})/\partial z_1$).

Offer functions for two firms, φ^1 and φ^2, are also depicted in Fig. 7.1. As indicated in the figure, these functions are convex in $\underline{z}$ as optimal offers increase at an increasing rate, holding profits and all else constant. Offer functions farther away from the horizontal axis represent higher levels of total profit. The firm's optimal choice of what level of z_1 to produce in each of its H units of the differentiated product is where the firm reaches the highest possible offer function while still being able to participate in the market. For Firm 2, this occurs at quantity z_1^{2*} and a total offer price of $P(z_1^{2*})$.

Figure 7.1 illustrates that the hedonic price function is simply an envelope of the equilibrium interactions between all buyers and sellers of a differentiated good. As such, it is possible for the hedonic price function to take any shape. The key point is that with relatively small data requirements, such as information on product types and their sales prices, we can recover the marginal implicit prices for any component characteristic of Z. The marginal price is equal to the revealed marginal WTP for that characteristic by consumers, $\partial\theta/\partial z_i$. In the Maine lakes example, Eq. (7.1) indicates that consumers' value for a one-unit (1 foot) increase in a property's footage of lake frontage is \$83 ($\partial\theta/\partial\mathrm{FRONT} = \partial P/\partial\mathrm{FRONT}$). This is the fundamental insight underlying the hedonic method. Researchers seek to estimate the parameters of the hedonic price function so they can recover information about the marginal value consumers place on characteristics.

[2] Another property of the bid function is that optimal bids increase proportionally with income $\partial\theta/\partial y = 1$. If income increases by \$1, the consumer's optimal bid must increase by the same amount to hold utility constant.

7.2 The Hedonic Price Function: Estimation

This section discusses the details of estimating a hedonic price function as well as some important threats to validity that researchers should consider as they attempt to uncover unbiased marginal value estimates. Table 7.1 summarizes these steps and considerations. As indicated, the first step is to gather the appropriate data; the basics of this task are discussed in Sect. 7.2.1. Common econometric specifications for the hedonic price function are discussed next in Sect. 7.2.1.1, which is necessary to ground the discussion of sample frame, data quality, and econometric methods that address data shortcomings that are presented in the remainder of Sect. 7.2. Before discussing each of the steps for estimating a hedonic price function, this section begins with a concrete example of a hedonic price function for housing that is based on a study by Boyle et al. (1999). Some of the study's results that were not reported in the journal article are also included in this chapter.

Boyle et al. (1999) estimated hedonic price functions for lakefront properties in Maine. Sales prices of properties, mostly summer cottages, were analyzed as a function of the characteristics of the structure on the property and important characteristics of the land, such as its location relative to the nearest town and the water quality of the lake on which the property is located. Lakes in Maine are known for their high water quality; however, this quality is being compromised in many Maine lakes by eutrophication resulting from nonpoint pollution. The physical manifestation of eutrophication is reduced water clarity. Thus, water clarity is the measure of lake water quality that is used by the authors in the hedonic price function.

Boyle et al. (1999) estimated a hedonic price function for each of four geographically distinct markets. The estimated hedonic price function, $P(\underline{z})$, for one market is:

$$\begin{aligned} P = 25,899 + 6,790 \cdot \ln(\text{SQFT}) + 83 \times \text{FRONT} - 3,919 \times \text{DNSTY} + 1,516 \times \text{DIST} \\ + 11,572 \times \text{HEAT} + 23,465 \times \text{BATH} - 17,579 \times \text{LKWATER} + 2.057 \times \text{WQ}, \end{aligned} \tag{7.4}$$

where the dependent variable is the sales price of a property, and the independent variables are, respectively, square feet of the structure on the property (mean = 750 square feet), the property's frontage on the lake (mean = 143 feet), the number of lots per 1,000 feet of frontage adjacent to the property (mean = 8.9 lots), the distance from the property to the nearest town (mean = 9.4 miles), and dummy variables that designate whether or not the structure has central heating, a full bath, or if the property uses lake water as its primary source of water. The last independent variable in Eq. (7.4) is a measure of water quality (WQ), which is equal to the area of the lake on which the property is located (mean = 4,756 a) multiplied by the depth of water clarity for the lake (mean = 3.9 m).

Table 7.1 Summary of steps for estimating the hedonic price function

Step 1	Collect data on property values and characteristics (Sects. 7.2.1 and 7.2.2)
	• Sales price: preferred measure of value, may need to consider selection bias • Pay careful attention environmental amenity/disamenity measurement • Develop appropriate neighborhood and locational variables • Develop a geographic information systems (GIS) database linked to census and environmental data • Consider appropriate sample frame for data (both across time and space)
Step 2	Choose functional form for the hedonic price function (Sect. 7.2.1.1)
	• Simple linear function in price and all characteristics generally not appropriate • Semilog functional form often used. Newer research suggests more flexible functional forms combined with spatial controls outperform simpler semilog form • Researcher judgment must be applied, and expectations about relationships between certain characteristics and sales price will guide choice of functional form
Step 3	Consider endogenous regressors potential due to omitted variables (Sect. 7.2.3)
	• Omitted variable concerns typically centered on omitted spatially varying characteristics • Spatial error model appropriate if omitted variables are thought to be independent of regressors • Spatial lag model and quasi-experimental designs can alleviate omitted variable bias
Step 4	Compute welfare measures (Sect. 7.3)
	• For marginal or nonmarginal changes in an amenity that are localized within a part of the market, the change in sales price resulting from the change in the amenity is the measure of net benefits if there are no transactions costs associated with moving between properties. If there are transactions costs, the change in price net of transactions costs measures net benefits (or is an upper bound on net benefits) • If a quasi-experimental design is used to estimate the change in sales price resulting from a nonmarginal change in an amenity, resultant capitalization rates are not likely to equal to net benefits • For nonlocalized changes in amenities, a second-stage demand analysis or sorting model approach is most appropriate for computing net-benefits

7.2.1 Data and Estimation Basics

At the most basic level, the data needs for estimating the hedonic price function and associated implicit prices are fairly simple. First, as indicated in Table 7.1, the appropriate dependent variable is sales price because the goal is to estimate the equilibrium price schedule, $P(\underline{z})$, which is a function that relates transaction prices (sales prices) to the characteristics of a product.[3] As with any durable asset, the sales price of a property represents the discounted present value (PV) of all future rents (R) from the property:

[3]Note: Alternatives to transactions prices, such as appraised value and owner-reported values, are also commonly used.

$$\mathrm{PV} = \sum_{t=1}^{T} \frac{E[R_t]}{(1+r)^t}, \tag{7.5}$$

where r is the discount rate, and T is the expected life of the property. As Eq. (7.5) makes clear, expected changes in future benefits associated with the house would be incorporated into current sales prices in a discounted fashion. Similarly, one can think of the implicit price for an individual characteristic as representing the current value of the discounted expected stream of benefits from that characteristic.

In some cases, the researcher might be interested in understanding how rental markets respond to changes in environmental quality (Grainger 2012; Baranzini and Schaerer 2011; Taylor and Smith 2000). In this case, the dependent variable is the contracted rental price and implicit prices are actually implicit rents, representing the additional value to rent from an additional unit of a particular characteristic. Rental prices are typically monthly rents but can be weekly rental prices in the case of vacation rentals. It is important to note future changes in amenities are not expected to be capitalized into current rents (as Eq. (7.5) makes clear). Although this does not diminish the usefulness of rental prices for hedonic analysis, the researcher has to be clear when interpreting implicit prices, noting that they represent the value of existing amenities in the current time period rather than changes in asset values (see also Epple et al., 2013).

Given a set of transaction prices, the next step is to identify the set of characteristics, $\underset{\sim}{z}$, that describe the good and influence its sales price. For example, the lakefront cottages used by Boyle et al. (1999) were assumed to have eight primary characteristics that determined price (as shown in Eq. (7.4)). Researchers must use their knowledge of the market to determine what characteristics are relevant for determining price in that market. However, there are important rules of thumb that apply to all circumstances. First, only those product characteristics that affect sales prices in a meaningful way are to be included as regressors. For example, the color of a product type often does not affect its price. Second, characteristics of the buyer and seller of the product *do not* belong in the hedonic price regression when markets are competitive (see also Taylor, 2008). This is not to say that characteristics of neighbors surrounding a home cannot be included—they can and should be included because they describe the neighborhood in which a home is located.

Three categories of regressors that are common to all hedonic housing studies are those that describe (1) characteristics of the house and its lot, (2) features of the home's neighborhood, and (3) the property's locational characteristics. There is no easy answer to the question of exactly what variables must be included within these three categories, but each is typically represented. For instance, housing features typically include at least the structure's size as measured by square footage or the number of bedrooms and baths. The age of the structure is usually included. Quality indicators, such as the presence of a fireplace, are sometimes included when available. The lot size is an important characteristic that is usually available and included in the analysis. Characteristics of lots are sometimes included, such as the amount of tree cover (Sander et al. 2010) or the view of mountains or water bodies afforded by the property (Bin et al. 2008; Baranzini and Schaerer 2011). Equation (7.4) indicates that Boyle et al.'s (1999) hedonic price function for summer vacation homes in Maine included square footage of the home, whether or not the summer home had central heat and a full bath, and whether it had lake water as its source of indoor water.

Neighborhood characteristics describe the immediate surroundings of a home and often include measures of the quality of the schools to which the home is associated and information about the general characteristics of the home's neighbors, such as median household income or racial composition. Neighborhood characteristics may also include an environmental amenity, such as ambient air quality, or undeveloped open space as a percentage of total land area in the neighborhood as a whole. In the Maine summer home example (Eq. (7.4)), the water quality of the lake on which a home was located and the density of development surrounding the home in question represented the key neighborhood characteristics.

Locational characteristics are conceptually similar to neighborhood characteristics in that they focus on a home's spatial relationship to other features of interest. In theory, locational characteristics vary by house, such as the distance of a house to an airport. In many cases, researchers will have access to very precise information on where a home is located and will be able to compute the exact distance of a home to features of interest. Boyle et al. (1999) included the distance of each summer home to the nearest town, and this varied by home (see Eq. (7.4)). However, in some cases the researcher may only know the neighborhood in which a home is located. Measuring locational characteristics at a coarser level, such as proximity of a neighborhood's center point to an airport, may be sufficient to capture the meaningful variation of the regressor. In this case, the distinction between a neighborhood feature and a locational feature is purely semantic.

There are many locational features that have been the focal point of hedonic studies using housing markets. Recent applications include studies that have focused on a home's proximity to amenities such as parks (Liu et al. 2013), oceans and water bodies (Landry and Hindsley 2011), and other forms of general open space (see McConnell and Walls, 2005, and Brander and Koetse, 2011, for recent reviews). The value of proximity to urban amenities, such as central business districts, transit hubs, or shopping centers has also been explored (e.g., Hess and Almeida 2007; Matthews and Turnbull 2007). Proximity to disamenities that have been studied include proximity to environmentally contaminated properties (e.g., Taylor et al. 2016; Zabel and Guignet 2012; Gamper-Rabindran and Timmins 2013; Kiel and Williams 2007), landfills (Hite et al. 2001), hog farms (Kim and Goldsmith 2009), power plants (Davis 2011), nuclear waste shipment routes (Gawande et al. 2013), and wind farms (Heintzelman and Tuttle 2012), among many others. The previous examples describe proximity of a property to a feature of interest, which conveys some form of exposure to the amenity or disamenity. Ambient conditions of a property have also been studied extensively, such as air quality (e.g., Grainger 2012; Chay and Greenstone 2005; Smith and Huang 1995) and noise reductions (e.g., Baranzini and Ramirez 2005; Pope 2008; Boes and Nüesch 2011).

There are a number of practical questions that an empirical investigator faces. What variables within each category described should be included in the hedonic equation? Should as many variables as possible be included for each category within the hedonic equation? These questions must be thoughtfully addressed. In an empirical investigation, the researcher should gather information from knowledgeable individuals (e.g., Realtors), review related hedonic studies, and test the robustness of his or her results to assumptions made during the modeling exercise. To help guide the researcher in his or her thinking, however, the goal should be remembered: an unbiased estimate of the implicit price of a particular characteristic. While inclusion of extraneous regressors may result in multicollinearity issues, this is a question of efficiency, not bias. The biggest threat to estimation of unbiased implicit prices is omitted variables. The potential for omitted variable bias is especially important in housing studies that focus on a locational characteristic given the multitude of locational features that may and that may or may not be known to the researcher. Given the importance of locational variables in nonmarket valuation, their measurement and how inclusion or omission of spatial features influences the researcher's ability to identify the implicit prices of any one particular locational feature of interest are discussed in detail in Sects. 7.2.2.2 and 7.2.3, respectively.

7.2.1.1 Functional Form of the Hedonic Price Function

In addition to specifying which variables belong in the hedonic regression, researchers must determine how the characteristics influence price. In other words, the researcher must decide on the functional form of the hedonic price function. Little theoretical guidance exists for the choice of functional form because the price schedule is determined in the marketplace by the interactions between many different buyers and sellers. If a product can be repackaged at no cost, the total price of the product would simply be the sum of the implicit prices of its component characteristics.[4] In this case, the appropriate specification for the hedonic price function would be a simple linear function:

$$P = \alpha_0 + \sum_{i=1}^{h} \beta_i H_i + \sum_{j=1}^{n} \beta_j N_j + \sum_{k=1}^{L} \beta_k L_k + \varepsilon, \qquad (7.6)$$

[4]The common example here is to think of a full grocery cart as "the product" that has component characteristics that are the grocery items contained within. In this case, there is no cost to repackage the component characteristics of the grocery cart, and so the total price for the product is simply the sum of the prices for each of its component characteristics.

where P is the sales price of a house; H represents structural and property characteristics of the house, such as square footage of the living area and lot size; N represents neighborhood characteristics, such as median income and quality of schools, and could include amenities, such as ambient air quality; and L represents locational characteristics, such as proximity to the central business district, and could include proximity to environmental amenities and disamenities. In the case of the linear functional form, the implicit price for any specific characteristic, z_i, is simply the estimated coefficient for that variable ($\beta_i = \partial P(\underset{\sim}{z})/\partial z_i$). With a linear hedonic price function, the incremental price of a property induced by an incremental increase in a characteristic is constant across all levels of each characteristic. This implies, for example, that the incremental value of an additional foot of lakeshore frontage would be the same for lakefront properties with 50 feet of frontage as for houses with 150 feet of frontage.

Of course, marginal prices are not likely to be constant for all characteristics, and so alternative functional forms are often used where some or all of the variables are transformed. For instance, most researchers assume that a home's price is related nonlinearly to its square footage. In particular, it is generally assumed that the total price of a house increases at a decreasing rate with square footage (see Fig. 7.1 and its associated discussion in Sect. 7.1). To capture this empirically, price may be estimated as a function of the natural log of square footage as in the Boyle et al. (1999) Maine summer home example (Eq. (7.4)), or price may be estimated as a function of square feet and a quadratic term for square feet.

A hedonic price function that allows for sales price to be affected nonlinearly by the characteristics of the property may be specified in many possible ways. Table 7.2 presents some common functional forms that introduce nonlinearity in the hedonic price function through transformation of the dependent variable, independent variables, or both. Also shown for each functional form is the specific form for computing the marginal implicit price, $\partial P(\underset{\sim}{z})/\partial z_i$. As can be seen from

Table 7.2 Functional Forms for the Hedonic Price Function

Name	Equation	Implicit prices
Linear	$P = \alpha_0 + \sum \beta_i z_i$	$\frac{\partial P}{\partial z_i} = \beta_i$
Semilog	$\ln\ P = \alpha_0 + \sum \beta_i z_i$	$\frac{\partial P}{\partial z_i} = \beta_i \times P$
Double-log	$\ln\ P = \alpha_0 + \sum \beta_i \ln\ z_i$	$\frac{\partial P}{\partial z_i} = \beta_i \times P/z_i$
Quadratic	$P = \alpha_0 + \sum_{i=1}^{N} \beta_i z_i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \delta_{ij} z_i z_j$	$\frac{\partial P}{\partial z_i} = \beta_i + \frac{1}{2}\sum_{j \neq i} \delta_{ij} z_j + \delta_{ii} z_i$
Quadratic box-cox*	$P^{\theta} = \alpha + \sum_{i=1}^{N} \beta_i z_i^{(\lambda)} + \frac{1}{2}\sum_{i,j=1}^{N} \delta_{ij} z_i^{(\lambda)} z_j^{(\lambda)}$	$\frac{\partial P}{\partial z_i} = \left(\beta_i z_i^{\lambda-1} + \sum_{j=1}^{N} \delta_{ij} z_i^{\lambda-1} z_j^{(\lambda)}\right) P^{1-\theta}$

*The transformation function is $P^{(\theta)} = (P^{\theta} - 1)/\theta$ for $\theta \neq 0$, and $P^{(\theta)} = \ln(P)$ for $\theta = 0$. The same transformation applies to lambda. The marginal effect is computed for $\theta \neq 0$. The marginal effect when $\theta = 0$ and $\lambda = 0$ is the same as for the double-log. The quadratic box-cox is equivalent to the linear box-cox, another common functional form when $\delta_{ij} = 0$ for all i and j. The marginal price of the linear box-cox is simply $\beta_i z_i^{\lambda-1} P^{I-\theta}$

these examples, the formulas can be complicated and can involve not only the estimated coefficients from the hedonic price function, but also the levels of the characteristics themselves.

Two possible approaches may be used for computing implicit prices when the formula involves the level of a variable. Consider the semilog functional form in which $\partial P(\underset{\sim}{z})/\partial z_i = \beta_i\, P$. In this case, the implicit price must be evaluated at some level of sales price. Most commonly, the mean or median housing price of the sample, over which the coefficients were estimated, is chosen. Alternatively, the implicit price may be computed for each house in the sample, and then the mean of these prices may be used. No one way is correct, although they have different interpretations. The former is the effect for the properties with the mean or median sales price, while the latter is the average effect for all sold properties. Ultimately, it is researcher judgment guided by the application at hand that will determine the appropriate level of price (or any variable) at which one evaluates the implicit price.

Product characteristics are often described by categorical variables. For instance, rather than recording the number of fireplaces in a home, a variable may be created to equal one if there are any fireplaces in the home and equal to zero if there are none. If the dependent variable is the natural log of sales price, then the interpretation of the coefficient estimate is the approximate percentage change in price when the characteristic in question is present. It is an approximation because the coefficients estimated for the dummy variables are transformations of the percentage effect (Halvorsen and Palmquist 1980; Kennedy 1981). For a coefficient estimate, b, the percentage effect, g, is given by $100 \times g = 100(e^b - 1)$. When b is small, the error in interpretation (without adjustment) is small. For instance, a coefficient estimate of 0.10 implies that the percentage change in price is actually 10.5%.

What can serve as a guide for choice in functional form? First, one should be careful not to rely on a linear hedonic (with no transformed variables) unless there are compelling logical reasons to do so. In general, nonlinear relationships between price and some size and quality attributes are expected. While researchers may use standard regression diagnostics to evaluate the fit of their chosen functional form, a classic paper by Cropper et al. (1988) takes a more rigorous approach to examining how choice of functional form may affect the researcher's ability to recover unbiased implicit prices. The authors conduct an exercise in which a simulation model of a housing market is calibrated with actual transactions data, and the resulting simulated equilibrium is used to compute the *true* marginal values for housing characteristics. These true marginal values are then compared to implicit prices computed from hedonic regressions estimated under two scenarios: one in which all housing characteristics are observed by the researcher, and one in which some variables are assumed to be unobserved by the researcher and thus are omitted from the hedonic regression. Each functional form presented in Table 7.2 was then evaluated for its ability to recover unbiased estimates of marginal values for specific characteristics. The authors found that simpler functional forms (e.g., the semilog) were better at recovering marginal values in the presence of unobserved housing characteristics. The findings of Cropper et al. (1988) have no doubt had a large impact on hedonic housing regressions over the past 25 years, as evidenced by the preponderance of semilog functional form choices made by authors, often with direct citation of Cropper et al.

More recently, Kuminoff et al. (2010) revisited the approach taken by Cropper et al. (1988) using current data and empirical methodologies, especially as relating to spatial features of housing markets. Contrary to Cropper et al., more flexible, functional forms such as the quadratic box cox are found to significantly outperform the simpler linear, semilog, and log-log specifications. The work of Kuminoff et al. has very important implications for current hedonic research, especially as researchers now consistently focus attention on controlling for unobserved spatial features of housing markets and often employ quasi-experimental methods as are described in Sect. 7.2.3.

However, note that box-cox estimation has been relatively uncommon within the hedonic housing literature. Early examples are Cheshire and Sheppard (1995), Palmquist and Israngkura (1999) and Ihlanfeldt and Taylor 2004. Given the implications of Kuminoff et al. (2010), researchers moving forward should expand their attention on flexible functional forms specifying the hedonic price function (e.g., Buck et al. 2014).

7.2.1.2 Market Definition: Space and Time

When choosing a sample of properties for a hedonic analysis, one must consider the geographic coverage of the data selected as well as the time period. First, consider the geographic coverage of the data. If a study's focus is an environmental good, then the data has to have geographic coverage sufficient to ensure variation in the environmental amenity or disamenity across properties. Depending on the amenity, variation may be in the form of the proximity of each property to the amenity (e.g., proximity to a park), or the variation may be the ambient level of the amenity (e.g., the level of air quality at the site). While variation of the first type is typically easy to ensure, ambient environmental quality can sometimes be more difficult to ensure in a sample frame. For example, in order to have sufficient variation in an environmental variable, the geographic dispersion of properties in a sample may be increased so much that properties are now drawn from different markets. In this case, estimating one hedonic price function for the entire sample is inappropriate because the hedonic price function is an equilibrium function describing a single unified market. The question that arises is how to determine what set of properties constitutes a single market. In other words, we wish to know the extent of the market.

Markets are truly separate if participants in one market do not consider houses in the other market when making purchase decisions, and sellers do not consider sales prices in the other market when negotiating sales prices. One common assumption is that each urban area represents a separate housing market. If the focus of the study is proximity to an amenity or disamenity, each urban area typically has sufficient variation for estimation.[5]

Although considering each urban area a separate market is likely a reasonable assumption, recent work has often assumed, either implicitly or explicitly, that housing markets are national (e.g., Hanna 2007; Noonan et al. 2007; Greenstone and Gallagher 2008; Grainger 2012; Sanders 2012).

In early work, Linneman (1980) considered whether housing markets are indeed national by comparing city-specific hedonic price functions for housing with a hedonic price function estimated on a national sample of housing. His hypothesis of a national market is neither fully supported nor rejected (early work is reviewed by Palmquist, 1991). The evidence for clear market segmentation across any barrier is mixed across applications, and it has often been the case that researchers do not provide evidence that supports their implicit choices.

It is difficult to test conclusively for market segmentation, but there are commonly employed methods that are used as supporting (or refuting) evidence for the researcher's priors about segmentation. The most common approach is to apply F-tests to determine if the hedonic price functions differ across segments (e.g., Taylor and Smith 2000). The problem with F-tests, however, is that results indicating that market segmentation exists may be due to mis-specification of the hedonic price function and not actual segmentation. F-tests are also likely to reject aggregation with large samples (Ohta and Griliches 1976). Once again, there are no definitive answers. However, researcher judgment along with supporting statistical tests can be used as a guide for determining the market extent in a particular study.

In addition to needing sufficient geographic variation in the data, it may be important to have appropriate time variation in the data. Often, a researcher is interested in how the valuation of an amenity changes over time. Or the researcher is interested in how prices changed in response to the creation of an amenity or disamenity, such as the siting of a landfill near a neighborhood. In some instances, the researcher may simply need observations from multiple years in order to increase sample sizes or because the research design and data dictate pooling data over many years.

When drawing a sample frame from multiple years, it is important to remember that the goal is to recover the parameters of a single equilibrium function that relates sales prices to a product's characteristics. As such, the parameters of the function should be constant over the time frame used in estimation. A single equilibrium hedonic price surface contrasts rather sharply with one's intuition about market conditions during the period 1996 to 2006 when U.S. *real* home prices had increased 86% during the decade and annual price increases of more than 10% were not uncommon in major cities around the world (Shiller 2007). However, the housing bubble burst in 2006, resulting in dramatic decreases in prices and

[5]Segmentation within an urban area has also been considered (e.g., Goodman and Thibodeau 2007).

increases in vacancy rates. Between 2006 and 2009, the U.S. housing stock was estimated to have lost nearly $4.5 trillion in value (Carson and Dastrup 2013).

Housing markets, like all markets, have cycles, and the issue is not so much one of appropriateness of using hedonic methods in cyclical markets, but rather careful understanding of how market cycles influence both estimation approaches and interpretation of results. Indeed, the dramatic consequences of the housing downturn in the late 2000s refocused the attention of hedonic researchers on how housing cycles affect best practices, especially given that modern transactions data available to researchers usually span a decade or more of sales (Boyle et al. 2012). The key question for nonmarket valuation is whether the changes in the hedonic surface are simply inflationary (or deflationary) and are easily netted out of the price function, or if housing market cycles impact the value of some housing characteristics differently than others.

If there are purely inflationary trends in a market, prices may be deflated prior to estimation by an appropriate, preferably local, price index. If an appropriate price index is not available, a series of dummy variables may be included in the hedonic regression to control for the year or quarter in which each property was sold. Either of these methods essentially control for changes in the intercept of the hedonic price function, and the assumption is that all implicit prices are changing by the same factor, and thus, adjustments to the hedonic price function intercept are all that is needed.

If underlying supply and demand conditions are changing over time, implicit prices of characteristics may change, and these changes may vary by characteristic (Cohen et al. 2014). Simple price deflators are not sufficient in this case, and the strategy taken may depend on whether there is upward or downward pressure on prices. During periods when prices are increasing, allowing for time period interactions with housing characteristics permits implicit prices to change over time and can be an appropriate strategy for capturing changing marginal implicit prices.

During cold markets, when there is downward pressure on prices, one must consider whether or not the housing market is in disequilibrium. The unique features of housing markets can lead to sticky prices after demand for housing shifts inward (Leamer 2007). In other words, excess supplies in housing markets, as indicated by higher than normal vacancy rates, can clear slowly as sellers are unwilling to adjust prices downward. Simple time-interaction variables can capture changing implicit prices during cold markets that are not severe, but if a market has substantial excess supply, as evidenced by high vacancy and foreclosure rates, additional modeling is needed to properly identify marginal prices.[6] As Coulson and Zabel (2013) discussed, both buyer heterogeneity and the composition of the stock for sale may differ in cold markets relative to normal periods. While these differences do not in and of themselves invalidate the approach of using time

[6]Note: Evidence suggests that foreclosures and vacancies can have a direct impact on neighboring property values (e.g., Campbell et al. 2011), and thus, proximity to these types of homes should be included as a disamenity in the hedonic price function to avoid omitted variables bias.

interactions to trace out changes in implicit prices over time, the problem is with interpretation of the marginal price as a representative marginal value. Zabel (2013) suggested that best practices should involve estimating the hedonic price function over a full housing cycle and then averaging the estimated implicit prices across the time period to compute an average implicit price for use in policy analysis.

How might one test for stability in the hedonic function? While F-tests may be applied to test for the stability of the hedonic function and the marginal prices of individual characteristics over time, the same weaknesses in the tests applies as stated previously for tests of market segmentation. Ohta and Griliches (1976) suggested a less stringent rule in which the standard errors for the constrained and unconstrained regressions are compared, and aggregation over time is rejected if errors increase by more than 10% (Palmquist 2006). As noted, substantial excess supply can be indicated by unusually high vacancy rates, and similarly, excess demand is measured by vacancies below the natural rate. The researcher should understand the market dynamics during the sample frame of the data and be aware that relying on data from housing markets that are making large adjustments is not likely to yield representative results for the implicit price of an environmental amenity (Zabel 2013).

7.2.2 Data Sources, Quality, and Measurement

A perfect dataset—ready for analysis and designed to answer a specific question that a researcher is interested in exploring—does not exist. The researcher will need to obtain data from multiple sources, each with their own challenges, and integrate them into a single database suitable for the analytical tasks at hand. This section explores a few commonly used types of data and highlights key challenges presented by the data for estimation of hedonic price functions.

7.2.2.1 House Value and Characteristics

Housing sales prices and characteristics are available from many sources. In many areas, the local tax assessor's office maintains electronic databases of their tax rolls that include a unique identifier for each parcel, its most recent sales price, lot size, a description of improvements (buildings) on the parcel, and other useful information, such as the school district to which the parcel is assigned (see Phaneuf et al., 2013, and Zabel and Guignet, 2012, for examples of microdata obtained from tax assessors).[7] These records are maintained at the county, city, or some other level of

[7]It can be the case that one residential property is formed by more than one parcel, and this is often true with commercial properties. The researcher must be careful that the final data used for analysis properly represents all the land (and buildings) contained in a single sale and that there is only one observation in the data for each unique sale.

municipal jurisdiction. Private vendors also collect local real-estate transactions data for re-sale in electronic formats (see Cohen et al., 2014, Kuminoff and Pope, 2013, Bajari et al., 2012, and Taylor et al., 2016, for references to sample vendors). While these data are often quite rich, they must be purchased. This contrasts with tax assessor records that are typically available free of charge or for nominal fees. For example, the King County tax assessor's office, which encompasses Seattle, Washington, provides rich data on sales prices, property characteristics, and locational characteristics via free download from its website.[8]

Data purchased from private vendors can have advantages over that obtained from tax assessors. First, privately provided sales data often includes all sales of each property over the time period for which data are provided. This can allow repeat sales models to be estimated, which is not possible when public data records only the most recent sale price for a property (e.g., Cohen et al. 2014). Second, assessor data may not be updated for property characteristics, whereas private data is often based on the National Association of Realtors multiple listing service data, which should be current. A shortcoming of private vendors is that data is often collected for metropolitan areas only. Tax assessor data are available for every location, including rural areas, although the quality and format in which the data are available may vary greatly.

Whether purchased from a private vendor or obtained directly from a municipality, records of actual transactions prices coupled with property characteristics are the norm in hedonic research. Once obtained, the researcher should carefully assess the data's quality.

First, sales prices may be recorded in error, resulting in unusually high or low sales prices. Even if sales prices are recorded without error, not all sales are arms-length transactions. In some cases, properties may be transferred to family members or from one business entity to another at prices much less than market value. These observations would not be appropriate for inclusion in the data set because the hedonic price function is an equilibrium function arising from competitive market bidding for each property. Unfortunately, recognizing transactions that were not arms-length is not always easy. An intuitive approach is to omit any observations on prices that are implausibly low or high. However, the researcher must carefully define what is too low or too high and should have a good understanding of the market.

Another consideration for sales data is the potential for sample selection bias. Here, the concern is that properties with specific unobservable characteristics may be less likely to be placed on the market for sale or less likely to sell once placed on the market. Sample selection bias may be particularly important in studies focusing on housing that is proximate to amenities or disamenities. For example, unobserved characteristics that influence the likelihood a home is placed on the market or is sold could be correlated with proximity of a house to a hazardous waste site or other disamenity. In this example, the ordinary least squares (OLS) estimate of the

[8]See http://info.kingcounty.gov/assessor/DataDownload/default.aspx; last accessed Nov. 5, 2013.

implicit price for distance from a hazardous waste site will be a biased estimate of the true market price, and the direction and magnitude of the bias will generally be unknown. It is not known if an explicit consideration of this issue has ever been undertaken. In a related paper, however, Huang and Palmquist (2001) developed a model in which they explicitly consider the simultaneity between the duration a property is on the market and its sales price. They argued that housing located near a disamenity may suffer two impacts: reduced sales prices and reduced probability of sales. They found that location near a noisy highway does not adversely affect duration on the market, but does adversely affect sales prices. However, they did not test for selection effects among the mix of properties offered for sale nor for bias in implicit price estimates, and this is an area where further research is warranted.

Finally, transactions data may be sparse and thus may need to be aggregated over a large number of years in order to have enough observations to reliably estimate implicit prices, especially as related to specific neighborhood or locational characteristics. However, aggregating over significant periods of time can create its own challenges, as was discussed in Sect. 7.2.1.2.

The use of survey data can overcome some of the limitations that transactions data may present. For instance, when research designs call for broad geographic coverage spanning multiple states and transactions data are prohibitively expensive to obtain, the researcher may use survey microdata collected by the U.S. Census Bureau.[9] Two types of microdata are available. The first is collected as part of the decennial census. In census years 2000 and earlier, the U.S. Census Bureau conducted a 1-in-6 sample of every household in the U.S. (approximately 18 million housing units in 2000), and these households were asked to answer what is referred to as the long form. The long-form survey asked a series of questions regarding housing that can be useful for hedonic researchers, including home value and basic housing characteristics.

Census survey microdata is publicly available. In other words, the exact survey responses for every individual who fills out a census form is made available to the public. However, the location of the household is confidential information, so the Census Bureau only releases survey results that identify the household's location by public use microdata area. Each public use microdata area is a statistical area that comprises at least 100,000 people.[10] Unfortunately, many, if not most, environmental amenities considered in hedonic housing studies vary at smaller scales than public use microdata areas, rendering the data incapable of estimating implicit prices for amenities and disamenities that vary substantially within a public use microdata area.[11] To address this limitation, some researchers are taking advantage of Research Data Centers through which the Census Bureau will grant researchers special access to the survey microdata that includes information on the city block in

[9]Microdata here refers to the individual survey responses.

[10]See www.census.gov/geo/reference/puma.html for more information.

[11]Bajari and Kahn (2005) used census public use microdata area in work analyzing racial segregation within cities.

which each home is located (see Gamper-Rabindran and Timmins, 2013, Liu et al., 2013, and Davis 2011).[12]

Finally, while individual survey responses linked to a home's location are not publicly available, the Census Bureau does publicly report aggregated housing statistics, such as median owner-reported housing value for census-defined neighborhoods. The smallest geographic area for which the Census Bureau releases median house values and housing characteristics is neighborhoods defined by "block groups," which can contain as few as 600 people but can contain up to 3,000 people. Block groups are meant to represent neighborhoods of similar people (see Noonan et al., 2007, for an example hedonic study using block groups).[13]

Researchers also use data aggregated by census tracts (e.g., Hanna 2007; Greenstone and Gallagher 2008). Census tracts are aggregations of block groups and usually contain between 1,600 and 8,000 people, although the Census Bureau indicates that 4,000 people is an optimum size for a tract. Census tract boundaries typically follow identifiable features (e.g., major roads) and do not cross county boundaries.

Census tract and block group summary data are attractive to researchers because of their ease of availability and universal coverage. However, given the aggregated nature of the data, the researcher has to pay close attention to the relationship between the spatial variation of the amenity of interest relative to the geographic scale of census block groups or tracts. When environmental amenities are expected to have external impacts that vary at a scale smaller than tracts or block groups, it can be difficult or impossible to identify their external effects. An excellent discussion of this point and how the researcher can address the problem within the context of publicly available census data is presented in Gamper-Rabindran and Timmins (2013).

A potential shortcoming of census data is that the Census Bureau discontinued the long form in 2000 and replaced it with the American Community Survey, which is a continuous survey that samples approximately 3 million addresses each year. A single sample does not fully cover the U.S. and is not likely to have enough observations in small geographic areas for valuing localized amenities. However, the Census Bureau samples in a way that the aggregate of samples over each contiguous five years has complete coverage of the U.S. and mimics the previous census long-form data in content.[14] The quality and appropriateness of this data for hedonic nonmarket valuation studies—especially in light of housing cycle impacts on implicit price estimates—is an area of research that has yet to be explored.

---

[12]See www.census.gov/ces/rdcresearch/ for more information.

[13]The smallest geographic area the Census Bureau defines is a census block, which is an area defined by streets or natural boundaries (e.g., imagine a city block in an urban area). Houses are identified by the block in which they are located only through special access to Census Bureau Research Data Centers. Census blocks are aggregated to block groups, usually containing approximately 40 blocks.

[14]See www.census.gov/acs/www/Downloads/handbooks/ACSGeneralHandbook.pdf for details on the American Community Survey. Last accessed Nov. 5, 2013.

An additional shortcoming of using census survey microdata is the potential for measurement error in the dependent variable because surveys ask individuals dwelling in a unit to estimate the value of the property. This is only a problem if the measurement error is systematically related to observed or unobserved characteristics of parcels. There are few studies that have systematically compared homeowner-assessed values to transactions data, but when they have, the results suggest little systematic bias in reported values.[15]

7.2.2.2 Neighborhood and Locational Characteristics

Most environmental amenities that researchers focus on in hedonic housing models are spatial in nature. Either the amenity directly varies over space (e.g., ambient air quality or density of tree cover) or the amenity is fixed in place and the researcher measures "exposure" of a home to the amenity through its spatial proximity (e.g., proximity to open space). Usually, measures of a home's exposure to a local amenity or disamenity have to be developed by the researcher directly. To do this, the first step is to develop a geographic information system database that contains accurate geospatial references for each property and each amenity of interest. Geospatial references are usually either boundaries of a feature, such as the boundary of a state park, or the center point of a feature, such as the center point of the central business district. Typically, the researcher will create a geographic information system database for properties in an area of interest and then merge this data with existing data obtained from municipalities, environmental agencies, or other entities that have data on environmental or locational characteristics of interest. By merging these two geographic information system data types together, the researcher creates a complete geospatial dataset with which to define (measure) a set of neighborhood and location characteristics for inclusion in estimation of the hedonic price function. Each of these steps is described more fully in turn, along with key considerations that can impact how an analysis is conducted.

If transaction microdata are collected from tax assessors' offices, one can usually also obtain a digitized tax map that contains the boundaries for every property in geographic information system format. These tax maps will identify each property by its unique parcel identification number, and that number can be linked back to the microdata recorded by the tax assessor, such as the most recent sales price of the property. If a digitized tax map is not available, street addresses can be used to

---

[15]Kiel and Zabel (1999) provided a comparison of owner assessments with actual sales prices for three metropolitan areas over approximately an 11-year period. Their results indicate that although owners generally overestimate the value of their homes by approximately 5%, the difference in owner valuation and actual sales price is not related to any housing or owner characteristic other than the tenure in the home (those living longer periods of time in their house at the time of the survey provided more accurate values). Although not a direct comparison, Gamper-Rabindran and Timmins (2013) compared overall results of an analysis conducted with national census data to parcel-level transactions data and find consistency across the two analyses.

assign latitude and longitude coordinates to the property using readily available software, such as ArcGIS. Address matching is not as precise as using a tax map because the location of a property is only approximate along the street segment in which the property is located.

Given a geographic information system database for properties, the spatial relationship between any one property and surrounding features of interest may be developed. Geospatial data for political boundaries (states, counties, municipalities) and transportation networks are available nationally from the U.S. Census Bureau (www.census.gov/geo/maps-data/), and they are also embedded within geographic information system software packages. Federal, state, and local agencies maintain rich datasets that are usually downloadable for free or for a very modest fee. Most university libraries also subscribe to geospatial data vendors that gather data from multiple sources and provide it in user-friendly formats. Luckily, rich geospatial data is typically available from urban county or city governments, planning agencies, and state and federal environmental management agencies. With a bit of searching, it is possible to build a rich geographic information system database of locational features for a study area.

An important source of neighborhood characteristics is the U.S. Census Bureau. Through the decennial census or American Community Survey, the Census Bureau publicly releases summary statistics of neighborhood characteristics, such as the percentage of owner-occupied homes and the racial composition and age distribution within a neighborhood. These data are released publicly at the census block group or tract level and are quite useful for measuring neighborhood characteristics.

The researcher can use block group or tract data in two ways. First, they can directly include summary statistics that capture neighborhood features of interest, such as median income or the percentage of owner-occupied homes in a neighborhood. Second, if the researcher is using sales price microdata, block group or tract-fixed effects can be included in the regression to capture all time-invariant characteristics associated with the home's neighborhood.[16]

Creating measures of environmental amenities is not necessarily a straightforward task. Ideally, one would have each characteristic measured in a manner consistent with the perceptions or understanding of the characteristic by the market participants. For instance, one can imagine many ways to construct a measure of proximity of a home to a nearby amenity. Linear distance of the center of a residential lot to a feature of interest is a common measure.[17] However, as Matthews and Turnbull (2007) indicated, measuring distance by the street network may be more important for highly localized amenities. In addition, proximity may not capture all the relevant features of "perceived exposure." For example, most studies

[16]When not using census tracts or block groups, researchers may use subdivision names, school districts, or local jurisdictions as fixed effects.

[17]Another common measure of proximity to a specific site or feature of interest is the use of buffer zones, such as a series of dummy variables that indicate if a property is within a certain distance band of a feature of interest (e.g., Taylor et al. 2016).

exploring the impact of hazardous waste sites on property values focus on the distance between a home and the nearest site.

Given agglomeration of land-use types, it is often the case that homes are proximate to multiple sites with varying intensity of environmental contamination (see also Leggett and Bockstael, 2000). The density of contaminated sites coupled with the intensity of contamination at each site is rarely addressed in studies. Furthermore, there may be several ways to measure the level of contamination present at a site, and it is unclear what measures best reflect market perceptions and, thus, are relevant to include in the hedonic price function.

If there is a high degree of correlation among measures relating to the same characteristic, including them all may lead to imprecisely estimated coefficients. On the other hand, dropping a variable to avoid multicollinearity could very well introduce bias. These are classical econometric issues. Clearly, balanced judgment on the part of the researcher, as well as sensitivity analyses on the choices made, are all part of the well-conducted hedonic study.

Measurement error in the constructed variables describing environmental amenities or disamenities is not limited to construction of spatial variables. For example, ambient environmental conditions such as air quality at the site may also be difficult to quantify in a manner that accurately reflects buyers' and sellers' perceptions of the characteristic. In Boyle et al. (1999), the Maine Department of Environmental Protection provided the measure of lake water clarity used in the hedonic regression analysis, and the clarity was measured monthly at a specific site on each lake during the summer months. Prospective purchasers may not have visited the lake during those months (and water clarity varies throughout the year on Maine lakes). As such, the objective or scientific measures of clarity collected by the Department of Environmental Protection may not have coincided with water clarity as perceived by prospective property purchasers. In a later analysis of the same market, Poor et al. (2001) found that an objective measure of water clarity at the time of sale was either preferred, or equally preferred, to the homeowner's subjective assessment for explaining variation in sale prices.

Although scientific measures or the proxy variables used by researchers are hoped to be reasonably well correlated with the *true* factors that are capitalized into a property's value, the use of proxy variables can nonetheless introduce the potential for errors-in-variables problems, leading to biased estimates of all coefficients. If only one variable is measured with error, improving the measurement of that variable will reduce the bias in all coefficients. However, when more than one variable is measured with error, reducing measurement error in one variable will not necessarily reduce the bias in the other regression coefficients (Garber and Klepper 1980).

Again, no definitive answer can be offered with respect to building the data needed for the hedonic model. Common sense, research and knowledge about the study market, and an understanding of the consequences of the choices (assumptions) made are key to conducting a thoughtful and careful hedonic price study. If multiple measures are available, the researcher should favor those that are consistent with economic intuition about what is likely to best represent the externality of interest as perceived by market participants, and report the sensitivity of the results to variable construction choices.

7.2.3 Endogenous Regressors

Endogeneity in hedonic models has long been a prominent research theme. Endogeneity can arise from simultaneous determination of sales price and a regressor or from omitted variables. In either case, the result is that coefficient estimates will be biased. Endogenous regressors are a classical econometric problem, and while the solutions follow classical techniques, there have been specific approaches focused on in the hedonic literature that are discussed here. The econometric approaches used to identify unbiased coefficient estimates are organized by the source of endongeneity: simultaneity or omitted variables.[18]

7.2.3.1 Simultaneity

First, consider endogeneity that arises from joint determination of sales price and a regressor. An example of this relates to land-use spillovers, which is particularly important for studies that seek to understand the value of proximity to open space for residential homeowners. Following Irwin and Bockstael (2001), consider the case in which all land is privately held and open space is potentially developable, as is typically the case at the urban fringe. For simplicity, imagine two properties adjacent to each other: parcels *i* and *j*. The value of parcel *i* in time *t* is a function of the amount of open space surrounding it—i.e., the open space on parcel *j*—and vice versa. However, the amount of open space around parcel *i* is an economic function of the value of its neighbor, parcel *j*, in development, and vice versa.

Addressing endogenous regressors that arise from joint determination relies on instrumental variables approaches (see Irwin and Bockstael, 2001, for a thorough discussion and empirical example). In other words, the researcher must find variables that are correlated with the endogenous variable (amount of open space surrounding a property in the above example) but uncorrelated with the error term. In considering open space around a particular property, Irwin and Bockstael used the characteristics of the land of neighboring properties (e.g., soil quality and land slope) as an instrument for the open space surrounding a particular property because these features serve as a proxy for the cost of developing the land.

[18]In this chapter and in the hedonic literature, two types of identification are discussed: identification of unbiased coefficient estimates and identification of the demand function for a characteristic of interest. Research that only estimates a first-stage hedonic price function would only discuss the former type of identification, while research that includes a second-stage analysis as presented in Sect. 7.4 is likely to include a discussion of both types of identification.

It is important to note that endogeneity due to the simultaneous determination of price and a characteristic of housing does not arise in a first-stage hedonic regression because housing prices are modeled as nonlinear functions of their characteristics. For instance, consider a commonly included characteristic—square footage of a home. Prices are usually modeled as varying nonlinearly with increased square footage. This makes intuitive sense because one would not expect an extra square foot of space to have the same marginal value for a 4,000-square-foot home as it does for an 800-square-foot home. However, square footage is not endogenous in a first-stage regression because of the fact that households that choose larger homes pay lower (or higher) marginal prices, ceteris paribus. The underlying parameters of the demand or supply of square footage is not being estimated in the first-stage regression, and as such, it is not endogenous. Of course, endogeneity of square footage could arise for other reasons. For example, square footage might be correlated with an important omitted variable, and if this is the case, the endogeneity would need to be addressed as described in the next section.

7.2.3.2 Omitted Variables: Spatial Econometrics and Quasi-Experiments

The second source of endogeneity being considered arises from omitted variables. As noted previously, housing location plays a critical role in most environmental valuation applications. As such, an important concern in hedonic housing research is that parameter estimates for spatially varying environmental amenities could be correlated with unobserved spatially varying characteristics.

Concerns over spatially varying omitted variables generally fall into two distinct categories. In the first case, one may believe that omitted spatial variables are independent of the regressors and only introduce spatial error correlation. In this case, spatial econometric methods may be introduced. Following Anselin (1988), a spatial error model may be written as follows:

$$\begin{aligned} P &= ZB + \varepsilon, \\ \varepsilon &= \lambda W \varepsilon + \mu, \\ \mu &\sim N\left(0, \sigma^2 I\right), \end{aligned} \tag{7.7}$$

where P is sales price, Z is a $N \times K$ matrix of property characteristics, B is a $K \times 1$ vector of coefficients, W is an $N \times N$ spatial weights matrix, ε is a $N \times 1$ spatial autoregressive error, μ is a $N \times 1$ random error term with variance σ^2, and λ is the coefficient on the spatially lagged error term ε. In the above, B and λ are estimated, and W is chosen by the researcher. The equations in (7.7) may be rewritten as:

$$P = ZB + (I - \lambda W)^{-1} \mu. \tag{7.8}$$

Specification of the spatial weights matrix, W, continues to be one of the more controversial aspects of spatial econometrics because results are sensitive to the choices made. The spatial weights matrix defines the sense in which properties are believed to be neighbors and determines the importance of any one observation to another. It is similar to a lag operator in the analysis of time series data but is multidimensional. In housing hedonics, distance-decay matrices are the most common specification of the weights matrix, wherein the importance of each property on the current property decays as distance increases. Alternative structures such as lattice matrices are also used. Lattices may be specified in several ways, such as allowing elements of the spatial weights matrix, w_{ij}, to equal one if a property shares a border with the observation of interest and equal to zero otherwise. Or, weights may be equal to one if properties are next to the property of interest or are contained within some predefined radius of the property of interest.

The second type of concern with regard to omitted variables is when they are thought to be correlated with other regressors, and classical endogeneity problems arise. In this case, parameter estimates will be both biased and inefficient. Two approaches are commonly employed to alleviate omitted variable bias. In the first, a spatial lag term (or spatial autoregressive term) may be added. Specifically, the model in Eq. (7.8) may be extended to allow for a spatial autoregressive process as follows:

$$P = \rho W_1 P + ZB + (I - \lambda W_2)^{-1}\mu, \tag{7.9}$$

where we now specify two weights matrices, W_1 and W_2, that are allowed to differ, and both ρ and λ are coefficients on the spatially lagged variables. If ρ and λ are estimated to be equal to zero, Eq. (7.9) reduces to a simple linear regression model. In a spatial autoregressive model, neighboring values are modeled as partially determining a home's value. This approach can be used to ameliorate the effect of omitted variables by using lagged prices as a proxy for unobserved neighborhood characteristics that contribute to the utility of owning a particular home in that neighborhood. The strength of this approach will rely on proper specification of W_1 and use of appropriate instruments for lagged prices because they too are endogenous.[19] Brady and Irwin (2011) provided an excellent, concise review of current best practices in applying spatial econometric models to common questions in environmental economics, while LeSage and Pace (2009) and Anselin (1988) provided technical introductions to spatial econometrics. Routines to implement spatial regression models are available in free statistical software, such as GeoDa or *R*, or proprietary software such as SpaceStat.

While the spatial autoregressive model can be a tool to recover unbiased parameter estimates for key variables of interest, there has been some confusion as

[19]While spatial autoregressive models can be employed to alleviate omitted variables, it should be noted that they may be used to directly model simultaneity as described in Sect. 7.2.1.1 (Irwin and Bockstael 2001).

to the economic intuition underlying the use of a spatial autoregressive model. Some have suggested that the intuition is that the spatial autoregressive model captures the price *determination* process rather than invoking a straightforward omitted variables argument. It is often suggested that because Realtors look to nearby property sales prices for "comps" (comparable homes) to help set the offer price for a home, these neighboring sales prices partially determine (influence) the sales price of a particular home. However, this is a mischaracterization of the price determination process. It is true that Realtors and homeowners use sales prices of nearby similar homes as an information tool to discover the nature of the equilibrium price surface and use this information to develop an initial offer price. If one believes the assumptions of the hedonic model, the ultimate sales price is determined by the interactions of many buyers with many sellers of competing homes in a competitive market. If a homeowner or Realtor misinterprets the information contained in nearby sales transactions and sets an offer price too far above the equilibrium for a home with its specific set of characteristics, the home will go unsold unless the owners are willing to reduce the price.

Note that the above argument does not indicate that a spatial autoregressive model should not be used. In fact, it can be an important means by which researchers can ameliorate potential biases in parameter estimates of interest that arise from omitted variables. The importance of the above point is highlighted when considering whether a spatial multiplier should be used to compute the benefits of an amenity improvement.

Briefly, a spatial multiplier approach indicates two benefits should be added together. The first is the direct benefit that an amenity improvement has on a property as measured by the marginal implicit price discussed in Sect. 7.1. To this direct effect, an indirect effect is added that is the impact of the amenity change on the price of the home in question through the channel of the autoregressive term in Eq. (7.9), $\rho W_1 P$. This latter effect occurs because the amenity change will also increase the price of neighboring homes (see Kim et al., 2003). However, as Small and Steimetz (2012) discussed, this indirect effect is only appropriate to apply if the *utility* of the homeowner is directly affected by the price of neighboring homes. In other words, if homeowners suffer from a type of money illusion and utility is impacted by the sales price of nearby homes irrespective of the real characteristics that underlie the prices, then a spatial multiplier is appropriate to add to benefits estimates.[20] This sort of pecuniary motivation seems unlikely.

[20]Another way to see this issue is to consider the reason why, or why not, lagged sales prices should be in the hedonic price function to begin with. This point is best explained with a simple example. Assume Home A is for sale. Home B is located near A, and was sold recently. Home B's characteristics may affect the utility of prospective buyers of Home A (e.g., attractive landscaping and home style). Inclusion of a lagged price after controlling for *all* real characteristics of nearby properties that affect the homeowner's enjoyment of his or her own home indicates that he or she receives utility just from the fact that a nearby home sold for a certain price. A pecuniary motivation for paying a higher price for Home A in this case indicates that people suffer from a type of money illusion or have speculative motives, neither of which is appropriate to include when calculating benefits of amenity improvements.

A second approach to alleviate omitted variables bias that has received growing attention in the hedonic literature is the use of quasi-experimental research designs. A quasi-experimental design broadly refers to a study design that exploits a transparent, exogenous source of variation in the explanatory variable(s) of policy interest that divides the properties into "treated" and "untreated" types. An example of treated properties could be homes that are located next to a hazardous waste site that is slated for cleanup due to a government program.

In the parlance of program evaluation literature, the evaluation problem is that researchers observe households in one state—treated or not—but never both (Moffit, 1991, and Meyer, 1995, provided concise introductions to program evaluation). To identify the effect of a treatment on the treated observations (e.g., the effect of hazardous waste removal on the price of nearby homes), valid counterfactuals for the treated observations need to be constructed. Researchers need to determine what the outcome of the treated group would have been had they not been treated. In classic experiments, researchers randomly assign individuals to treatment and thus can expect the outcomes of the control group—had they actually been treated—to be equivalent to the outcomes of the treated group up to sampling error. Through randomization, the control group is expected to provide a valid counterfactual for the treated group—they provide an estimate of the outcome of the treated group had they *not* been treated.

In hedonic modeling, the researcher is generally not able to randomize houses to a treatment. Instead, the researcher must rely on a naturally occurring process, such as a natural event or a policy decision that assigns homes to treatment in an exogenous manner (i.e., in a way that is unrelated to the error term in the hedonic regression). For this reason, quasi-experiments are also referred to as natural experiments. For example, Petrie and Taylor (2007) exploited a natural experiment to estimate the value of access to irrigation for agricultural land. They exploited a policy decision made by state regulators who issued a (surprise) halt to the issuance of irrigation permits in a particular watershed. Irrigation permits had previously been essentially free and available to all landowners that applied. After the policy change, the only way to acquire a new permit was to purchase land that had already been granted a permit (permits were not transferable otherwise). The policy change thus turned a free asset into a scarce asset whose value should be reflected in sales prices of agricultural land with permits. It was reasonably argued that the policy applied to agricultural parcels in a way that was unrelated to factors that determine land prices, and properties just outside the policy watershed boundary served as a control group. The authors could then exploit both time (pre- and postpolicy change) and space (inside and outside the watershed) to identify the value of access to irrigation by estimating the implicit price of permits in a hedonic regression model.

When exploiting a natural experiment to identify the value of an amenity, the process by which individuals (or households in this case) are selected into treatment defines the tools needed to identify a treatment effect. If selection is solely determined by observable characteristics, then matching estimators may be employed (Todd, 2008, provided an excellent overview of matching estimators). While this approach has received some attention in the hedonic housing literature (see, for example, Liu and Lynch, 2011), it has been more common to employ methods that address selection based on unobservables.

As discussed previously, a primary concern has been unobserved spatial characteristics that might correlate with amenities of interest. If there is selection into treatment based on unobservable differences between the treatment and control groups, estimated treatment effects will be biased. The two main approaches for alleviating this problem are to employ difference-in-difference designs or regression-discontinuity designs. Parmeter and Pope (2013) provide an excellent introduction to quasi-experimental research designs as applied in hedonic housing research.

To describe the most common difference-in-difference model in a hedonic housing context, consider a situation in which the researcher has available repeated cross-sectional data, and there is a policy that randomly assigns some homes to the treatment status. Following Parmeter and Pope (2013), a difference-in-difference model in this case can be written as

$$\ln(P_{ijt}) = \alpha_0 + \alpha_1 d_t + \alpha_2 d_j + \beta_0 d_{jt} + \sum_{k=1}^{K} \beta_k z_{ki} + \varepsilon_{ijt}, \tag{7.10}$$

where i denotes the observation, j denotes the group ($d_j = 1$ if treated, $d_j = 0$ if not treated), t denotes the time period ($d_t = 1$ after the policy change, $d_t = 0$ before the policy change). Thus, $d_{jt} = 1$ if the property is treated and the sale is after the policy change; otherwise, $d_{jt} = 0$. All other characteristics that determine sales prices are included in the vector z. In Eq. (7.10), pure time effects are captured by α_1, and pure group effects are captured by α_2. The effect of the treatment on the value of the treated houses (those affected by the policy) is given by β_0. To see this, consider the example above and say that a researcher has a sample of all home sales in a market within one-half mile of a contaminated property. In this case, $d_j = 1$ if a home is within one-half mile of an environmentally contaminated property that receives cleanup, and β_0 captures the increase in house price due to cleanup of a nearby contaminated site—the treatment effect of interest.

Note: For reasons discussed in Sect. 7.3, this treatment effect is referred to as the capitalization rate (Kuminoff and Pope 2014). The parameter α_1 captures any time-trend differences pre- and post-policy across the entire sample, while α_2 captures any time-invariant differences between housing surrounding sites that are cleaned up versus those surrounding sites that are not cleaned up.

A related approach to a difference-in-difference model is the regression-discontinuity design that can be employed when the researcher has a single cross-section available. In a regression-discontinuity design, assignment to treatment is determined fully or partly by the value of a covariate relative to a threshold. In the simplest formulation, there is a deterministic function of a single covariate known as the forcing variable, D, where $D_{1i} = 1$ if and only if $z_i \geq c$. Here, all z_i with a value of at least c are assigned $D_1 = 1$. This sharp design gives rise to a model identical to that in (7.10), except that rather than time, t, there is a different boundary. For example, the model may be written as

$$\ln(P_{ijs}) = \alpha_0 + \alpha_1 d_s + \alpha_2 d_j + \beta_0 d_{js} + \sum_{k=1}^{K} \beta_k z_{ki} + \varepsilon_{ijs}, \quad (7.11)$$

where all is as defined in Eq. (7.10) except that s denotes a spatial boundary ($d_s = 1$ if a home lies within the boundary, and $d_s = 0$ otherwise). In this instance, one exploits discrete, exogenous variation over space. In a regression-discontinuity design, one limits the sample to those influenced by the forcing variable (close to the boundary) to alleviate omitted variable bias in the s and j dimension. For instance, in a non-environmental context, homes on either side of a school attendance boundary could be used to estimate the value homeowners place on school quality (see Black, 1999).

In hedonic applications, the above models exploit discontinuities in time and/or space to identify treatment effects. Several types of quasi/natural experiments have been exploited in this context. One type takes advantage of sharp changes in information available to homebuyers. Changes in seller disclosure laws were used by Pope (2008) to identify the value of reducing airport noise. Davis (2004) used a sudden, unexplained increase in cancer cases in a community (and the associated news coverage) to value reductions in cancer risks. These studies highlight how changes in governmental information about local conditions can provide opportunities for the researcher to examine how locational disamenities are capitalized into housing prices.

A number of studies have used natural events such as hurricanes to identify the value of changes in flood risk perceptions (e.g., Hallstrom and Smith 2005; Bin and Landry 2013) or a nuclear transit accident to identify the value of avoiding risks associated with being close to a transportation route (Gawande et al. 2013). Perhaps the most common type of application in hedonics has been to take advantage of policy changes that induce quasi-random assignment to treatment. Here, treatment status—such as being located near a hazardous waste site that is remediated or being located in a watershed where new irrigation permits are no longer available—is determined by known factors that serve as valid instruments. A few examples include Chay and Greenstone (2005), who used changes in the Clean Air Act to value air quality; Petrie and Taylor (2007), who used a change in a state's regulatory permitting process to value access water for agricultural production; and Greenstone and Gallagher (2008) and Gamper-Rabindran and Timmins (2013), who valued hazardous waste site cleanup by exploiting particular features of the rules used by the U.S. Environmental Protection Agency for determining what sites would receive funding for cleanup.

Although natural experiments are a useful strategy for addressing endogeneity concerns, they too have underlying assumptions and data vulnerabilities that do not make them a panacea for identifying unbiased treatment effects. First—and key among the conditions that must be met to identify treatment effects in all applications—is that selection into the treatment must not depend on unobservables. This condition can be hard to ensure in housing applications because it is often the case that broad samples (across time or space) are needed to gather enough housing sales observations for analysis. Other underlying assumptions that are critical to understand and test their validity will vary by approach and model specification. For instance, Eq. (7.10) assumes that implicit prices for all other characteristics, β_k, are equal across groups. If the β_k is not equal across j, then one might worry about unobservable characteristics that vary across j as well. In addition, one assumes the samples are selected from the same population in the same way over time and across groups. This might be violated if, say, the change in policy results in a different mix of housing being placed for sale postpolicy.

Another very important consideration for nonmarket valuation is that the situations and data that allow one to implement a quasi-experimental approach may result in estimated treatment effects that only reflect capitalization rates but do not reflect underlying willingness to pay for the change in the amenity (Kuminoff and Pope 2014). This point is discussed more fully in the next section that develops the set of welfare theoretic measures of net benefits from an amenity change that can be computed by estimating a hedonic price function.

7.3 Welfare Measurement with the Hedonic Price Function

Our model indicates that the implicit price of amenity $i(\rho_{zi} = \partial P(\underline{z})/\partial z_i)$ is equal to the consumer's marginal WTP for the amenity ($\theta z_i = \partial \theta(\underline{z}, y, U)/\partial z_i$). Implicit prices are the most commonly reported result from hedonic studies. If one is interested in whether or not a current stock of an amenity is capitalized into the market for housing, estimates of the marginal implicit prices (sign, magnitude, and statistical significance) are the appropriate measures to report. In the context of expenditures on other aspects of housing, these relative prices can provide interesting insights about the importance of various housing amenities. Of course, it should be recognized that if an amenity change affects consumers in other ways (say, increases the recreation benefits to consumers who do not own nearby property but visit the park after it is improved), these benefits would not be captured using the measures discussed in this chapter.

In instances where we want to know the value consumers might place on a *change* in an environmental amenity, the relationship between implicit prices and appropriate measures of WTP for the change depends on the situation. To examine this issue, this section considers owners of properties and renters of properties separately. Even though the owner and renter can be the same person (implicitly renting from oneself), this discussion considers these two sides of the market separately. Two types of changes are considered: a change in a localized amenity and a change in a nonlocalized amenity. Examples of localized externalities that have been studied in the hedonic context are highway noise (Palmquist 1992a), a hazardous waste site (Kohlhase 1991), and an incinerator (Kiel and McClain 1995). In these cases, a change in the amenity affects a relatively small number of properties, so the equilibrium hedonic price function for the entire market is unaffected. When changes in nonlocalized amenities, such as the air quality of a city occur (Zabel and Kiel 2000; Sieg et al. 2004), the amenity change affects a large enough number of houses to reflect a marketwide shift in supply, and thus, one would expect a change in the market-clearing equilibrium hedonic price function.

For a marginal change in an amenity is localized, first consider the effects on renters of the property if no transaction costs are associated with moving. When a decrease of an amenity occurs, the renters are no longer at their optimal bundle of housing given that they face the same hedonic price schedule as before the change at their home. If renters can move without cost to a house with the original bundle of characteristics, there will be no change in welfare for the renters. However, the owners of the property realize a capital loss on the property because of the decrease in the amenity level associated with the property. The owner would be willing to pay an amount of money up to the value loss of the property value to avoid the amenity change. If, in addition to being a localized change, the amenity change is marginal (i.e., a one-unit change), then willingness to pay is simply the implicit price, $\partial P(\underline{z})/\partial z_i$, for each property owner. The total willingness to pay is the sum of the implicit prices across property owners who receive a change in the amenity.

If the amenity change is localized and nonmarginal in the magnitude of change, owners would be willing to pay an amount equal to $P^1 - P^0$, where P represents the sales price of the property with the initial level of the amenity (P^0) and the new level of the amenity (P^1). The total WTP is the sum of the price changes for houses that receive the amenity change. If the analysis is done prior to an environmental change taking place, as is typical for policy analyses, the price change is forecast using the estimated hedonic price function.[21] Thus, for a change in an amenity, say characteristic z_1, the total change in welfare or net benefit (NB) is computed in the following way:

$$\mathrm{NB} = \sum_{k=1}^{N} P_k\left(z_{1k}^1, z_{2k}^0, \ldots, z_{nk}^0\right) - P_k\left(z_{1k}^0, z_{2k}^0, \ldots, z_{nk}^0\right), \tag{7.12}$$

where $P_k(z_{1k}, z_{2k}, \ldots, z_{nk})$ is the hedonic price function evaluated with the kth property's characteristics in either the initial or new state.

For example, using the estimated hedonic price function in Eq. (7.1), a property with characteristics equal to the sample means (see discussion following Eq. (7.1)

[21]In the case of a localized amenity, researchers are able to forecast the change in price prior to the amenity change actually occurring because the hedonic equilibrium does not change. One still must be careful that the magnitude in the change of the amenity is not outside the range observed in the total sample. If so, the prediction is out of sample and becomes less reliable the further out of sample the change.

for those means) is predicted to sell for \$110,466. If the water clarity at the lake, where the property is located (a small part of the overall market) were increased from 3.9 to 6.3 m (the highest observation for this data), this property would be predicted to sell for \$115,133. The net benefit for this house is \$4,667. The total net benefit would be the sum of the net benefits estimated for each house located on this particular lake.

If one relaxes the assumption of costless moving on the part of renters, then the computation above represents an upper bound on the welfare change associated with the amenity change (i.e., overstate gains and understate losses). From the price change given in Eq. (7.12), one would need to subtract the transactions costs (TC) associated with each renter having to move to his or her new optimal location in order to compute the total change in welfare. For each property, net benefits are given by $P^1 - P^0 - \mathrm{TC}$. Palmquist (1992b) summarized the manner in which transactions and moving costs may be quantified and incorporated empirically into a hedonic analysis.

The above analysis assumes consumer utility is unchanged as a result of the change in the amenity (Δz_1) because renters move to a new house with the exact same bundle of characteristics as before the amenity change. This follows from the assumption that the change in the amenity does not affect the overall hedonic price function in the market but only affects the price of houses experiencing Δz_1. Thus, in optimizing utility, a house identical to the original dwelling before the amenity change will be chosen. If identical housing is not available and utility changes as a result of the move, then the correct welfare measure would require quantifying the difference in utility arising from the consumer no longer being at the optimal consumption bundle (see Bartik, 1988, for a complete description of this problem). While this computation is not possible, using information from only the hedonic price function, the change in rents less transactions costs would provide an upper bound on the resulting net benefits from an amenity change. The degree to which this bound is close to the true net benefits will depend on the degree of substitutability between the housing affected by the amenity change and the amenity level at the housing renters move to.

Table 7.3 summarizes the benefit estimates possible with the first-stage hedonic price function when the change in the amenity is localized. Note that the amenity change that occurs locally may be marginal or nonmarginal. The table highlights that simply computing the change in property values at locations will overstate the benefits of an amenity improvement and will understate the losses of an amenity decline when transactions costs are associated with moving (Bartik 1986).

When transaction costs prohibit moving, the net benefits of an amenity change cannot be exactly measured, but an upper bound can again be computed as the price differential for the property minus the transaction costs associated with moving. Consider a decline in an amenity. The landlord's capital loss as a result of the amenity change is exactly offset by the tenant's capital gain associated with paying a lower rent (recall that the tenant does not move). Thus, the price change is a transfer between owner and tenant, and no net benefit is associated with the change in rent. However, the reduction in rent only partially offsets the consumer's utility

Table 7.3 Benefits calculations with the hedonic price function

Computation	Appropriate situation
Cross-sectional estimators	
$\partial P/\partial z_i$	Net benefits of an ↑/↓ in the amenity by one-unit (a marginal change) in a localized context with zero transactions costs associated with moving
$P^1 - P^0$	Net benefits of an ↑/↓ in the amenity in a localized context with zero transactions costs associated with moving
$P^1 - P^0 - TC$	Net benefits of an ↑/↓ in the amenity in a localized context when transactions costs are positive, but do not prohibit moving. Household can relocate to an identical house as before the amenity change
$P^1 - P^0 - TC$	*Upper-bound* on net benefits of an ↑/↓ in an amenity in a localized context when (1) moving is possible, but households cannot relocate to an identical house as before the amenity change or (2) when it is *not* possible to move because transactions costs are prohibitively high
Quasi-experimental estimators	
ΔP	The treatment effect, commonly measured by $\partial P/\partial d_{jt}$ (see Eq. 7.10), is measure of net benefits if change is marginal and localized, but unlikely to equal net benefits for nonmarginal changes (even if localized)

loss. Transaction costs provide an upper bound on the remaining utility loss (after the change in rent is taken into account). Thus, the difference between the change in rent and transaction costs associated with moving will provide an upper bound on the total welfare change associated with the change in the amenity for a consumer (Table 7.3). Palmquist (1992b) presented a graphical exposition of this point.

In the case of *nonlocalized* amenities, an exact measure of welfare change cannot be computed with only information from the hedonic price function. This is because a nonlocalized change implies a shift in the overall supply, and the hedonic equilibrium itself will change. The new hedonic price function cannot be forecast using information from only the current price function. Bartik (1988) demonstrated that the change in hedonic rent at sites that are improved, as predicted by a hedonic price function estimated prior to a nonlocalized change in an amenity (call this measure ΔHP), is likely to be an upper bound for the net benefit of the amenity change. This bound is computed as given in Eq. (7.12) and is, therefore, quite easy to implement. However, ΔHP is only an upper bound under some conditions on the profit changes landlords may experience when allowed to make equilibrium adjustments to housing characteristics in response to the amenity change (Bartik 1988). These profit changes are likely to be in a range such that ΔHP will be an upper bound only if no transactions costs are associated with moving, which is likely implausible. If there are significant transactions costs, this measure is less likely to be an upper bound. Unfortunately, one cannot know if ΔHP is an upper bound, and empirical tests have not determined how likely it is to be an upper bound.

Note also, the computation of net benefits discussed above is directly linked to our understanding of what quasi-experimental designs measure (see Sect. 7.2.3.2). Recall, quasi-experimental estimates rely on an exogenous shock to the market—usually a nonmarginal shock—to identify changes in prices due to a change in an amenity. The treatment effect, or change in price due to a change in the amenity, is measured by β_0 in Eq. (7.10), and it is only equal to the underlying WTP for a change in the amenity under certain conditions. As Kuminoff and Pope (2014) showed, the shock to the environmental amenity must be uncorrelated with the initial conditions (i.e., the initial housing *and* environmental conditions) and changes in the housing stock over time. However, this may be problematic in practice because many environmental policies are by design correlated with initial environmental conditions. Further, it is hard to ensure that behavioral responses to changes in environmental amenities are not significant (e.g., a change in the composition of housing postenvironmental improvements). This last consideration can be alleviated by focusing the analysis on a relatively narrow time window around a particular shock, but data limitations may not always allow this either.

In general, the measure of ΔP (see Table 7.3) from a quasi-experiment will not equal net benefits when ΔP is measuring differences across two different hedonic equilibria (one prepolicy equilibrium and one postpolicy equilibrium) except in the limiting, unrealistic case when demand for an amenity is perfectly elastic.[22] Kuminoff and Pope (2014) also demonstrated that the bias in the measured ΔP (a "capitalization rate") will generally be in an unknown amount and direction, and need not be bounded by ex ante or ex post willingness to pay. Finally, even if the exogenous shock is randomized in a way that satisfies the conditions needed to identify marginal willingness to pay in the postpolicy equilibrium, caution is still needed in using postpolicy marginal willingness to pay to assess the benefits of a policy. For example, a postpolicy marginal willingness to pay of zero for a further improvement in environmental quality could still mean that the policy increased welfare substantially because prepolicy marginal willingness to pay was high, or hardly at all because prepolicy marginal willingness to pay was low (Kuminoff and Pope 2014).

7.4 Demand Analysis Within the Hedonic Framework

To estimate uncompensated demands for the *characteristics* of the differentiated good, information on the quantities of characteristics purchased and the marginal implicit prices of those characteristics from a hedonic analysis are combined with information on the socioeconomic characteristics of the purchasers. Estimating the demand for characteristics of a differentiated product using hedonic prices is often referred to as a second-stage analysis (estimating the hedonic price function was described as the first-stage in the previous section). While Rosen (1974) provided the framework for this method, it continues to be refined today. This section

[22]See also Klaiber and Smith (2013) who employed a simulation strategy to further explore the relationship between capitalization rates and willingness to pay for amenity changes.

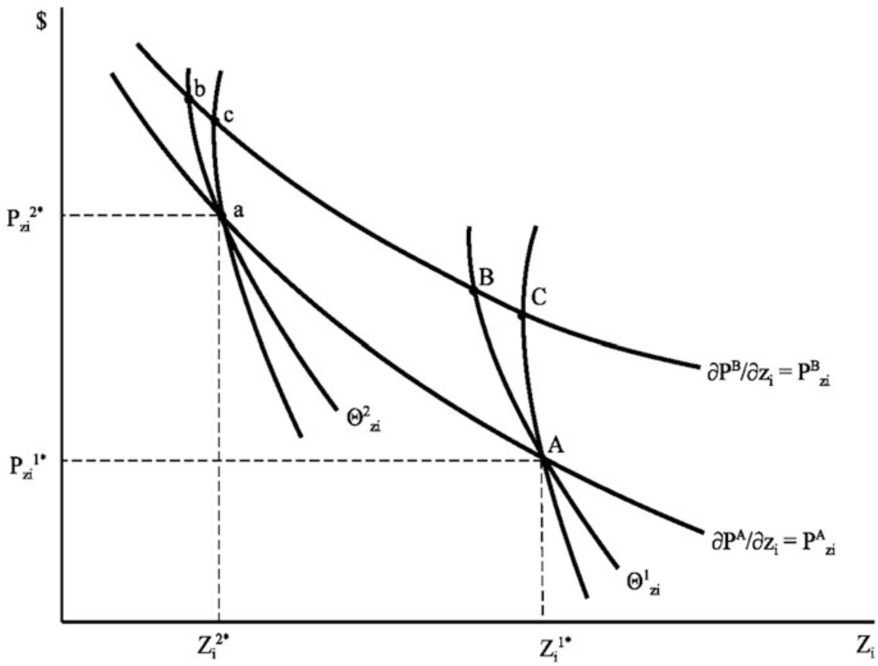

Fig. 7.2 Demand identification with the hedonic price function

provides an overview of second-stage methods. See Palmquist (2006) for a more technical discussion of the method.

For welfare measurement, researchers are typically interested in compensated demands, which indicate the willingness to pay for z_i, holding the levels of all other goods/characteristics and utility constant. Recall that the bid function for each individual, θ^j, is a function describing the maximum WTP for a specific bundle of characteristics, holding utility and income constant. Because bid functions change directly with income (i.e., $\partial\theta/\partial y = 1$), the marginal bid function for characteristic z_i, θ_{zi}, only depends on $\underline{z}$, utility, and exogenous consumer characteristics, α^j (Palmquist 2006). As such, the marginal bid function is equivalent to an inverse compensated demand function for characteristic z_i.[23] It describes the change in WTP for z_i as the level of z_i changes, holding utility and all other characteristics and goods constant.

Figure 7.2 illustrates the marginal bid function for two individuals, θ^1_{zi} and θ^2_{zi}. For the moment, consider only the marginal bid function for Consumer 1 as

[23]Note: The marginal bid function is not equivalent to a marginal rate of substitution function given by $(\partial U/\partial z_i)/(\partial U/\partial x)$, as had been suggested in the early literature. It is true that, at the *optimal level of consumption*, the marginal rate of substitution is equal to the marginal bid and the marginal price $(\partial U/\partial z_i)/(\partial U/\partial x) = \partial P/\partial z_i = \partial\theta/\partial z_i$, as discussed in Sect. 7.1.

depicted by the heavy dark line passing through points *AB* and the associated implicit price function P^A_{zi}. We can see that the marginal bid function is equal to the implicit price function at the optimal level of consumption for Consumer 1, z_i^{1*} (Point *A*). This point corresponds in Fig. 7.1 to the optimal consumption level where the bid function is tangent to the hedonic price function.

Theoretically correct measures of welfare change for a characteristic of a differentiated good are no different than was described for a quantity change in an amenity in Chap. 2. The literature typically makes the distinction between measures of welfare change when consumers face a price change (compensating and equivalent variation) and when consumers face an exogenous quantity change (compensating and equivalent surplus). These are analogous measures (see Chap. 2). In the case of an exogenous change for consumers who cannot move, compensating and equivalent surplus would be the appropriate measures. In the case of an exogenous change in z_i in which moving is possible, compensating and equivalent variation would be more appropriate. However, these measures will only measure the benefits to homeowners. If the change in the amenity affects consumers in other ways (say, increases recreation benefits to consumers who do not own nearby property but visit the park after it is improved), these benefits will not be captured using the measures discussed in this chapter.

Here, the focus is on compensating or equivalent surplus, as one typically assumes households cannot move in response to a change and thus computes the households' willingness to pay as a lower bound to the true welfare change (see Sect. 7.3). Measures of compensating or equivalent surplus are simply computed by integrating under the compensated inverse demand or the marginal bid function:

$$W(\Delta z_i) = \int_{z_{10}}^{z_{11}} \frac{\partial \theta(z_i; \underline{z}, U^j)}{\partial z_i} \mathrm{d}z_i, \tag{7.13}$$

where $W(\Delta z_1)$ is compensating or equivalent surplus for a change in z_i depending on whether U^j is equal to the initial level of utility or the level of utility after the amenity change, respectively. This measure of welfare change is appropriate for the situation in which consumers receive an exogenous quantity change in an amenity, and consumers are compensated in terms of the numeraire (holding z constant). Palmquist (2006) discussed welfare measures when other types of compensation are used to return consumers to their original (or new) level of utility.

Of course, estimating the marginal bid function directly is not possible because the marginal bid depends on unobserved utility. However, several alternatives allow us to recover the information necessary to compute welfare measures.

The first approach follows our discussion to this point and models the consumer's choice among a smooth continuum of housing bundles in which all possible combinations of characteristics are possible. In this approach, a second-stage analysis is used to recover Hicksian uncompensated demand if data from multiple markets are available.

The second approach relies on random utility modeling, which refocuses the modeling exercise to describe how a consumer makes a discrete choice to purchase a single house from a set of alternatives that he or she is considering. In this approach, utility parameters are directly estimated and welfare analysis follows directly. Each method is described in turn next.

7.4.1 Uncompensated Demands for Characteristics

Uncompensated demand functions can be derived analytically in a manner analogous to a homogenous goods market if the hedonic price function is linear so that the marginal prices are constant. However, researchers often expect the hedonic price function to be nonlinear and the marginal prices to be nonconstant. In this instance, the budget constraint is no longer an exogenous constraint; the marginal price of any characteristic is simultaneously determined by the choice of the amount of that characteristic to purchase. Traditional optimization methods are inappropriate in this case.

Uncompensated demands with nonconstant marginal prices can be derived, however, if we linearize the budget constraint around the optimal consumption bundle of the consumer. Palmquist (1988) demonstrated that utility maximization with this constraint set yields the same choices as would be observed with the original, nonlinear budget constraint. The linear budget constraint is of the form

$$Y_a^j = x + \sum_{i=1}^{n} P_i^* z_i, \tag{7.14}$$

where Y_a^j is Consumer j's income adjusted in the following manner:

$$Y_a^j = y - P(Z*) + \sum_{i=1}^{n} p_i^* z_i^*. \tag{7.15}$$

The linearized budget constraint in (7.14) is exogenous as the implicit prices, p_i^*, faced by Consumer j are held constant at the level associated with this consumer's actual purchase, z_i^*. Income must be adjusted as shown in Eq. (7.15) because a linear budget constraint will imply a larger consumption set than is affordable for the consumer with income y facing a budget constraint in which prices for z_i are decreasing as amounts of z_i are increased.

The linearized budget constraint in (7.14) may be used in conjunction with the first-order condition for utility maximization (given in Eq. (7.3)) to solve for the inverse uncompensated demand in a manner analogous to that for homogeneous goods:

$$P_1^j = f\left(z_1, z_2, \ldots z_n, x, Y_a, \alpha^j\right). \tag{7.16}$$

This demand function can be estimated using quantities of characteristics chosen, the implicit prices of those characteristics, adjusted income, and other socioeconomic characteristics assumed to affect demand.

Consumer surplus estimates for a change in z_i may be computed by integrating vertically under the inverse demand curve for z_i between the initial and the new level of z_i (Parsons 1986). For example, suppose the following simple linear demand for z_1 is estimated (note, this is not an inverse demand function):

$$z_1 = \alpha + \beta_1 p_1 + \beta_2 p_2 + \delta y, \tag{7.17}$$

where p_2 is the implicit price of a substitute/complement for z_1 and y is adjusted income. The integral of the inverse demand between the initial level z_{10} and new level z_{11} is

$$\int_{z_{10}}^{z_{11}} \frac{1}{\beta_1}(z_1 - k)\mathrm{d}z_1 = \left[\frac{1}{\beta_1} z_1^2 - \frac{k}{\beta_1} z_1\right]\Bigg|_{z_{10}}^{z_{11}}, \tag{7.18}$$

where k is a constant equal to $\alpha + \beta_2 z_2 + \delta y$ evaluated at some appropriate level, such as the mean observed in the data.

The demand function estimated should include relevant characteristics that are substitutes or complements to the characteristic of interest. Choosing the appropriate set of substitutes and complements may be somewhat ad hoc because no theoretical guidance defines exactly what is a substitute or complement for any one characteristic. A reasonable approach is to assume that the major characteristics of a property (such as house, lot size, and a few key neighborhood characteristics) are likely substitutes and complements to include. Boyle et al. (1999) estimated the following uncompensated demand for water clarity (Q_{wc}):

$$Q_{\mathrm{WC}} = f\left(P_{\mathrm{WC}}, P_{\mathrm{SQFT}}, P_{\mathrm{FRONT}}, Y_{\mathrm{adj}}, \Gamma\right), \tag{7.19}$$

and computed consumer surplus for changes in water clarity as described in Eqs. (7.17) and (7.18). In Eq. (7.19), Y is income as adjusted in Eq. (7.15) and P_{WC}, P_{SQFT}, and P_{FRONT} are the marginal implicit prices of water clarity, square feet of living space, and lake frontage of the property, respectively. They chose two major characteristics of the property: size of the house and the amount of lake frontage. Results indicated that both characteristics were substitutes for water clarity as the price coefficients for house size and lake frontage were positive and significant.

In addition to income, the researcher must determine the other relevant factors expected to shift demand, represented by the vector Γ in Eq. (7.19). As with any demand function, typical factors might include the age and education of the purchaser. For property markets, factors such as the number of children under 18 years old would also be expected to be important demand shifters. Other factors specific to each application might be included. For example, in the Maine lakes example, Boyle et al. (1999) included a dummy variable indicating whether or not the owner had friends who had bought property at the lake prior to his or her own purchase of property.

After the set of substitutes, complements, and demand shifters are chosen, the researcher must still specify the function to be estimated. A common functional form for demand is the semilog functional form. Boyle et al. (1999) estimated a linear, semilog, and Cobb–Douglass specification for Eq. (7.19), preferring the nonlinear specifications. Based on the semilog specification, the estimated own-price elasticity for water clarity was −0.158, and the estimated surplus for an improvement in lake water clarity from 3.78 to 5.15 m was $3,765. The estimated surplus for the same change in quality was $12,870 using linear specifications and $3,677 using Cobb–Douglas specifications. As might be expected, the results are sensitive to the choice of functional form. A careful researcher investigates how sensitive their results are to choice of functional form (and specification of variables included). If they have a preferred equation for reporting, reasons for this preference should be explained.

Consumer surplus measures can be computed using estimates of the parameters of the uncompensated demand. However, consumer surplus is only an approximation of the theoretically correct measures of welfare change (see Chap. 2, and Freeman et al., 2014). Two approaches may be used to recover compensating or equivalent variation when uncompensated demands are estimated. In the first, utility is specified and a system of uncompensated demands are derived analytically and estimated. Given estimates of the demand parameters, duality may be used to analytically recover estimates of compensating or equivalent variation by solving for indirect utility or expenditure functions. This approach was taken by Palmquist and Israngkura (1999) and Parsons (1986).

Alternatively, one can estimate a single demand for a characteristic of interest and use differential equation methods to recover an associated utility or indirect utility function. Hausman (1981) and Vartia (1983) demonstrated how this may be accomplished when prices are changing. Palmquist (2006) demonstrated this method for quantity changes to solve for the bid function.

Consider the inverse uncompensated demand in Eq. (7.16). Recognizing that the marginal implicit price is equal to the marginal bid function at the optimal level of consumption, we can substitute $\partial\theta/\partial z_i$ for the left side of Eq. (7.16). Also, because our numeraire x is identically equal to $Y - \theta$, we can substitute for x in the right-hand side, and Eq. (7.16) becomes a differential equation in θ. Solving this differential equation for the bid function allows us to compute welfare measures as follows:

$$W(\Delta z_i) = \int_{z_i^0}^{z_i^1} \frac{\partial\theta(z_i, \underline{z}, U^j)}{\partial z_i} \mathrm{d}z_i = \int_{z_i^0}^{z_i^1} \mathrm{MRS}\left(z_i, \underline{z}, x\left(z_i, \underline{z}, U^j\right)\right), \tag{7.20}$$

which is equal to compensating surplus if utility is held constant at the original level of utility, and equal to equivalent surplus if utility is held constant at the level of utility after the change (note: income is held constant).[24]

7.4.1.1 Identification of Uncompensated Demand

Recall that each consumer will choose an optimal level of each characteristic by finding the level of consumption at which the marginal bid for z_i equals the implicit price of z_i. By estimating the parameters of $P(z)$ and computing $\partial P(\underline{z})/\partial z$, one can observe one point on each consumer's uncompensated demand as well as his or her compensated demand function or marginal bid function. (The term "marginal bid function," rather than inverse compensated demand function, is used for consistency with past discussions.) However, no other information on the shape of the bid function is observed. Consider Fig. 7.2 again, which shows the marginal bid functions for two consumers and the marginal price function P^A_{zi} (ignore for the moment the second marginal price function, P^B_{zi}). For reasons discussed in Sect. 7.2.1, the hedonic price function is likely to be nonlinear, so the marginal price function is likely to be nonconstant. In Fig. 7.2 the marginal price of z_i is shown to decrease as more of z_i is consumed.

Figure 7.2 indicates that each marginal price reveals only one point on each consumer's marginal bid function, so one cannot make inferences about the shape of the marginal bid function using only information on marginal prices. By observing P_{zi}, it is known that one point on the marginal bid function of Consumer 1 is $\{P^{1*}_{zi}, z^{1*}_i\}$; however, it cannot be determined if the demand function is shaped as shown by the solid line or the dashed line. Indeed, an infinite number of possible functions could pass through the point $\{P^{1*}_{zi}, z^{1*}_i\}$. This is the standard problem of identifying any demand function: Additional information besides the market-clearing prices for each property in a market is needed to identify demand.

One approach to identifying demand is to use information from separate markets, such as different cities that are geographically distinct. In this approach, one assumes that individuals in one market do not consider housing in the other markets as substitutes when making purchase decisions. Thus, the hedonic price function in each market is distinct and arises from interactions of only those buyers and sellers

[24] An alternative method for recovering the information necessary to compute compensating and equivalent variation is to directly specify the functional form for household utility and derive the functions to be estimated (see Quigley 1982). In other words, the exact specification of the regression is determined by the analytical solution to the utility maximization problem. What function is estimated depends on the goal of the research. For estimating welfare changes, Cropper et al. (1993) and Chattopadhyay (1999) estimated the equilibrium condition given in Eq. (7.3) in which marginal prices (computed from the first-stage hedonic regression) are equal to the marginal rate of substitution function. While this function may not be used directly for computing welfare changes, it recovers utility parameters that can then be used to compute estimates of welfare using duality results.

located in that market. Individuals with any given vector of characteristics are assumed to have preferences over attributes that are identical across markets. However, because of differences in supply, or perhaps because of the distribution of socioeconomic characteristics among individuals within a market, the equilibrium hedonic price functions will vary across markets, so similar people will be observed making different housing choices across the different markets. If this is the case, estimating separate hedonic price functions in each market will identify demand. This point is illustrated in Fig. 7.2, where P_{zi}^{B} represents the marginal price function from a separate market. Given this additional information, we can now determine if it is Point *B* or *C* that is the optimal choice for that type of consumer and whether the marginal bid function is represented by the dashed or the solid line.

Thus, to estimate an inverse uncompensated demand (such as in Eq. (7.16)) or a system of demands, hedonic price functions are estimated for multiple markets. The marginal prices from each market are pooled with information on property owners' demographics and characteristics of their properties to estimate demand. For example, to estimate the demand given in Eq. (7.19), Boyle et al. (1999) specified a hedonic price function like that given in Eq. (7.1) for four markets in Maine. Each market was distinguished by distances between each other, by unique regional characteristics, and by different real estate multiple listing service regions. The estimated marginal implicit prices along with corresponding quantities purchased were pooled with demographic information from a survey of property owners.

The use of multiple markets to identify demand for housing attributes has also been used by Palmquist (1984), Parsons (1986), Bartik (1987), Cheshire and Sheppard (1998), Palmquist and Israngkura (1999), and Zabel and Kiel (2000), with the latter two studies focused on environmental amenities. The number of markets used to identify demand has varied from two (Cheshire and Sheppard 1998; Bartik 1987) to seven (Palmquist 1984; Parsons 1986) to 13 (Palmquist and Israngkura 1999). There is no established correct minimum number of individual markets required. Of course, what must be established is that the hedonic price functions do vary across markets; thus, researchers must provide evidence that statistically significant differences in the implicit prices across markets exist. The issues associated with determining appropriate assumptions for market segmentation are the same here as are discussed in Sect. 7.2.1.2.

Kuminoff and Pope (2012) proposed an alternative strategy within a quasi-experimental setting. Specifically, the authors show that if a natural experiment meets the requirements to be a good instrument and provides a large shock so that the hedonic equilibrium shifts, then this known, exogenous shift in the price function may be used to identify the inverse uncompensated demands. This point can be illustrated by reviewing Fig. 7.2, which highlights how identification could be achieved with information from separate markets. In a quasi-experimental setting, one could estimate a single-period hedonic price function preshock (akin to P_{zi}^{A} in Fig. 7.2) and a separate single-period hedonic price function postshock (akin to P_{zi}^{B} in Fig. 7.2) to provide identification for the inverse bid functions.

7.4.1.2 Endogeneity of Prices and Income in Demand

In addition to identifying demand, an important econometric issue that must be taken into account is the possible endogeneity of implicit prices and income. Recall that for any functional form of the hedonic price function other than linear, as given in Eq. (7.6), implicit prices may be nonconstant. If prices are nonconstant, then they are endogenous because consumers simultaneously choose the marginal price they pay per unit of the characteristic when they choose the level of the characteristic. Also, in the case of nonconstant implicit prices, we linearized the budget constraint to analytically derive the demand function, which involved adjusting income by the implicit prices (Eqs. (7.14) and (7.15)). When including adjusted income in the demand specification, it too will be endogenous because the adjusted income relies on the nonconstant implicit prices.

For instance, the hedonic price function estimated by Boyle et al. (1999) in Eq. (7.1) indicates that the marginal price for water clarity and square feet of living space are nonconstant. Thus, by choosing the quantity of living area, for instance, the purchasers are also choosing the marginal price per foot of living area they must pay. The marginal price of living area and the choice of square feet of living area are both endogenous.

Endogeneity is typically handled by instrumental variable techniques. In this method, each price that is endogenous and adjusted income is first regressed on a set of exogenous explanatory variables. These exogenous variables or instruments should be related to the endogenous variable of interest but exogenous to the system. The instruments must be (1) correlated with the regressors, (2) uncorrelated with the error term, and (3) of full rank (that is, adds new information). The resulting parameter estimates from the first-stage regression are used to predict values for the endogenous prices and income. The predicted prices and predicted income are then used to estimate the demand equation (see Greene, 2011, for a standard discussion of two-stage least squares).

Choosing the proper instruments for the endogenous variables can be a difficult task. For the demand given in Eq. (7.19), Boyle et al. (1999) developed nine instruments for the marginal prices that described factors such as number of lakes in a market area, distance of the market to Boston, a time trend, and local economic and demographic conditions (e.g., an index of real estate activity in the area and the percentage of the current population that is employed). For adjusted income, the instruments were socioeconomic characteristics of the property purchasers (as reported in a survey of the owners) and included age and gender of the owner, number of children under 18-years old, educational level, retirement status, and the number of people in the household. Also included were the squared terms for age, the number of children under 18-years old, and the number of people in the household.

Palmquist (1984) estimated the demand for four housing characteristics; he used exogenous socioeconomic characteristics (age, marital status, number of dependents, race, and their square where appropriate), and he used a set of dummy variables for each of his seven urban areas to instrument for the nonconstant implicit prices and adjusted income.

Cheshire and Sheppard (1998) used spatially lagged variables as their instruments. The authors developed a spatial relationship between each house in their data set and each next-closest house.[25] For each observation, the characteristic prices and income levels associated with the two houses closest to the house in question were used as instruments. The authors argued that these variables clearly meet requirements (1) and (3) for instrumental variables that were discussed earlier and are likely to meet requirement (2) because there was enough variability in neighborhood housing characteristics (and individual characteristics within a neighborhood). Palmquist (2006) made the point that because this approach is not unidirectional (i.e., the spatial lags are not unidirectional in the same way time lags are), these instruments are invalid. To alleviate this problem, one might combine time and the spatial lags to create an instrument for the implicit price that is the closest house sold prior to the current house.

A common issue is that instruments do not explain the variation in the endogenous variables very well. In finite samples, weak instruments result in biased estimates of the demand parameters, just as in the case of using simple, ordinary least squares (Bound et al. 1995). Thus, with weak instruments, one may be accepting higher variance of the parameter estimates without reducing the bias. Also, if instruments are weak and the loss of efficiency is large, common tests for endogeneity (Hausman 1978) are likely to have weak power (Nakamura and Nakamura 1998). In addition, a loss of relevance can occur when the instruments chosen for the endogenous regressors do not adequately capture the type of variation in the regressor. For these reasons, Nakamura and Nakamura rejected the "always instrument" policy for endogeneity—especially for variables that are not of primary interest to the application or when it is not clear that the potential instruments are truly exogenous. While there is no clear guidance for what are weak instruments, R^2 values for the auxiliary equations of less than 0.2 caused concern for Nakamura and Nakamura.

7.4.2 Discrete-Choice Models of Housing Demand

A different approach to identifying demand parameters recasts consumer decisions in a random utility framework.[26] In a random utility model, consumers are assumed to make a discrete choice between house bundles, rather than a continuous choice over attribute levels as in the typical hedonic framework. In the random utility

[25]Closeness is defined by both geographic proximity and similarity of the structures.

[26]See Klaiber and Kuminoff (2014) and Kuminoff et al. (2013) for detailed reviews. Phaneuf and Requate (in press) also provided a general overview. Note, the model discussed here in the context of housing choice is conceptually and empirically equivalent to the random utility model described for travel cost models in Chap. 6.

model framework, the consumer knows his or her utility for all possible choices, but utility is measured with error because not all factors that influence choices are accessible to researchers. Thus, utility (*U*) is assumed to be the sum of a systematic portion (*V*) and a random component. This can be written as follows for a consumer, *j,* who is observed choosing a house, *k*:

$$U^j\left(x^k, \underline{z}^k; \alpha^j\right) = V^j\left(x^k, \underline{z}^k; \alpha^j\right) + \varepsilon_{jk}, \tag{7.21}$$

where $U(\cdot)$ is the true, but unobservable utility as defined in Eq. (7.2), and ε is the error term introduced because the researcher cannot observe all relevant aspects of the individual. The individual maximizes utility by selecting the house that yields the highest level of utility from among the set of all possible alternatives, *A*. The probability that Consumer *j* selects House *k* is given by

$$\begin{aligned} &\Pr(k|A) = \Pr(U_k \geq U_m) = \Pr\left\{V\left(x^k, \underline{z}^k; \alpha^j\right) + \varepsilon_{jk} \geq V\left(x^m, \underline{z}^m; \alpha^j\right) + \varepsilon_{jm}\right\}, \\ &\forall \;\; k, m \ni A \text{ and } k \neq m. \end{aligned} \tag{7.22}$$

The assumption about the distribution of the random error term implies which probabilistic model is estimated. A common assumption is that the error terms are independently and identically distributed following a Type 1 extreme value distribution, leading to the well-known conditional logit model. To relax the independence of irrelevant alternatives assumption inherent in the conditional logit model, a nested logit model or a random parameters logit may be used.

Welfare measurement for a marginal change in z_i is given by $(\partial V/\partial z_i)/(\partial V/\partial x)$, and for a nonmarginal change, the willingness to pay, *C*, for a change in an amenity of interest, say z_1, is given by $V\left(x^0, z_1^{k0}, z_2^k, \ldots z_n^k; \alpha^j\right) = V\left(x^0 - C, z_1^{k1}, z_2^k, \ldots z_n^k; \alpha^j\right)$. Note that in this formulation, the price paid for the house is held constant at the purchase price. Substitution of the budget constraint, $x^0 = y^j - P(\underline{Z}^k)$, makes this clear. In this sense, the welfare measures from the random utility model approach rely on the current distribution of prices. Further, because the random utility model in itself does not model the market equilibrium, it cannot be used to infer how a nonlocalized change in an amenity would affect the equilibrium price schedule. Thus, while the random utility model can directly compute welfare measures based on the current price vector, additional steps are needed to estimate general equilibrium welfare estimates for a nonlocalized amenity change (e.g., air quality changes across a city).

Sorting models extend the random utility model to address general equilibrium responses to a nonlocalized change in an amenity. Environmental applications of sorting models shift the focus of the analysis from the choice of a house to the choice of a neighborhood and add an equilibrium condition equating demand and supply for each neighborhood. Environmental amenities in these models vary across neighborhoods, and housing supply is typically assumed to be fixed. Aggregate demand is equilibrated to fixed supply by the equilibrium price vector for the n neighborhoods in the model,$\{P_1, \ldots P_n\}$.

Demand in a sorting model is defined by a household's expected demand for each location. Equation (7.22) can be recast to represent this demand by letting k denote neighborhoods rather than houses and having P and z be indexes of these variables computed by the researcher to vary by neighborhood (see Klaiber and Phaneuf, 2010, for a discussion). Substituting the budget constraint into Eq. (7.22) and summing over individual demands yields the expected aggregate demand for neighborhood k:

$$\mathrm{AD}_k(P, z, \alpha) = \sum_j \mathrm{Pr}_{jk}\left(P, z, \alpha^j\right), \tag{7.23}$$

where P and z are the vector of prices and housing characteristics across the entire landscape (i.e., for all k neighborhoods). Writing Eq. (7.23) in share form gives the proportion of market participants who are expected to select neighborhood k, given a set of prices and attributes across the landscape. We then define an equilibrium condition, which requires that the share of demand for any particular neighborhood, k, must be equal to the proportion of total housing in the market that is available in neighborhood k:

$$s_k^D(P, z, \alpha) = \frac{1}{J} AD_k(P, z, \alpha) = s_k^S. \tag{7.24}$$

The market-clearing condition in Eq. (7.24) is important for welfare analysis. Given a shock to z (say, an improvement in air quality), Eq. (7.24) can be solved for a new set of market-clearing prices postshock. As illustrated below, this allows computation of welfare effects from a change in z that do not require prices to be held at their original, preshock levels as traditional random utility models require.

In addition, a particularly nice feature of these models is the straightforward manner in which they allow the researcher to incorporate unobserved neighborhood attributes. To see this and how the mechanics of estimation proceed, rewrite utility in Eq. (7.21), focusing on the choice of a neighborhood:

$$U^j\left(P^k, \underline{z}^k, q^k, \zeta^k; \alpha^j\right) = \alpha P^k + \beta Z^k + \delta q^k + \gamma \alpha^j q^k + \zeta^k + \varepsilon_{jk}, \tag{7.25}$$

where utility now depends on an index of housing price in neighborhood k, P^k, housing characteristics in neighborhood k, Z^k, and an environmental attribute of interest in neighborhood k, q^k. Equation (7.25) allows for heterogeneity in preferences for the environmental amenity through the term $\gamma \alpha^j q^k$. Further, neighborhood unobservables are explicitly incorporated in the model through ζ^k, a measure of neighborhood attributes that is observable to the decision-maker but unobserved by the researcher.

Equation (7.25) may be rewritten in a simplified form that makes clear the role of attributes that vary by neighborhood and those that vary by individual:

$$U^j\left(P^k, \underline{z}^k, q^k, \zeta^k; \alpha^j\right) = \mu^k + \gamma\alpha^j q^k + \varepsilon_{jk}, \tag{7.26}$$

where

$$\mu^k = \eta P^k + \beta Z^k + \delta q^k + \zeta^k. \tag{7.27}$$

In Eq. (7.26), the choice of the kth neighborhood depends on the population mean utility for neighborhood k (average tastes), and $\gamma\alpha^j q^k$ differentiates utility away from the mean for households based on observable characteristics, α^j.

Estimation now proceeds in two steps. In the first step, Eq. (7.26) is estimated by assuming the error in Eq. (7.21) is distributed Type I, extreme value leading to the conditional logit model. The conditional logit model has the convenient property that the estimated demand share for any particular neighborhood will equal the observed share, and these shares will sum to one across all neighborhoods. In other words, the conditional logit model imposes the equilibrium condition on the data as an artifact of the model's statistical properties. Given the assumed utility in Eq. (7.26), this can be written formally as

$$s_k^D \equiv \frac{1}{J}\sum_j^J \left(\frac{\exp(\mu^k + \gamma\alpha^j q^k)}{\sum_m \exp(\mu^m + \gamma\alpha^j q^m)}\right) = s_k^S. \tag{7.28}$$

The maximum likelihood estimates of μ^k and γ assure Eq. (7.28) holds.

In the second step, the estimated μ^k are then used to estimate Eq. (7.27), which decomposes the μ^k into observable (P^k, Z^k, q^k) and unobservable (ζ^k) components. In this stage, endogeneity concerns surround the correlation between P^k and ζ^k. Given that the estimation of Eq. (7.27) is linear and estimated by ordinary least squares, standard econometric tools (instrumental variables) are available to address this endogeneity.

Given the full set of parameter estimates, general equilibrium welfare effects of a policy change can be evaluated. With a change in q^k, Eq. (7.28) is used to solve for the new price vector. Welfare measures that incorporate both changes in q and changes in prices are then available to compute based on the assumed form for utility. Klaiber and Phaneuf (2010) implemented this framework to explore the value of open space amenities. See also Tra (2010, 2013) for applications valuing air quality.

It is worth repeating that strategies for identifying unbiased coefficient estimates must be considered carefully in the random utility model framework as well. Even though random utility models and sorting models operate somewhat differently, they still require careful consideration of endogenous variables due to simultaneous determination or omitted variable bias just as with multimarket approaches. In addition, random utility models and sorting models require careful consideration about the definition of choice sets and the unit of analysis (a neighborhood).

7.5 Labor Markets and the Value of a Statistical Life

In addition to housing, a second important application of the hedonic method in nonmarket valuation is labor markets (Rosen 1974). To see how hedonic modeling applies within a labor market context, reconsider Fig. 7.1. In a labor market context, the hedonic price function is now a hedonic wage equation, and it is an envelope of individual firm's offer functions and workers' indifference curves. For example, Fig. 7.1 could be the hedonic wage gradient for a specific job characteristic, such as a measure of the on-the-job risk of injury. The wage trade-offs that workers are willing to make to accept a higher risk of on-the-job injury, holding utility constant, are represented by Φ. Higher contour levels of Φ would represent higher levels of utility. Similarly, θ are now iso-profit curves of the firm, representing the trade-offs the firm can make between higher wages and expenditures to reduce the risk of on-the-job injury, holding profit constant. Lower contour levels of θ represent higher levels of profits (see Bockstael and McConnell, 2007, for an excellent technical review).

At first blush, hedonic wage models may not seem directly related to environmental valuation. However, as Cameron (2010) highlighted, they continue to be the subject of vigorous academic and policy debates because they are a primary way in which one measures the benefits of federal environmental (and other) regulations that reduce human mortality risks. Specifically, hedonic models are used to estimate the trade-offs workers are willing to make between wages and the risk of death on the job. These dollar trade-offs can then be used to compute a measure referred to as the "value of a statistical life" or VSL. The VSL is not the value of saving a particular individual (an "identified life"), but the value society places on a reduction in the probability of one death among them (see Viscusi 1992, 1993).

The easiest way to understand the VSL is through an example. Consider the following simple wage hedonic:

$$\text{wage}_k = \alpha + \beta_r \text{risk}_k + \sum_{n=1}^{N} \lambda_n X_{kn} + \sum_{m=1}^{M} \gamma_m D_{km} + \varepsilon_k, \tag{7.29}$$

in which the wage of the kth worker is estimated to be a function of the risk of death on the job (risk_k); N variables describe human capital and demographic characteristics of the worker (X_{kn}), such as age and education; and M job characteristics (D_{km}) other than the risk of death, such as whether or not supervisory activities are associated with the job.[27] In a linear specification such as (7.29), the implicit wage for risk, $\partial w/\partial r = \beta_r$, is the additional wages a worker would require to assume an additional increment of risk of death on the job. By normalizing over the risk

[27]The compensating wage equation describes the equilibrium wage as a function of the factors affecting the wage negotiation. This negotiation will be in regard to factors such as job duties and working conditions, as well as characteristics of the workers themselves that affect expected productivity.

increment, the compensating wage differential for risk is converted to the value of a statistical life. For example, suppose risk is measured in units of deaths per 10,000 workers, wages are hourly earnings, and a simple linear compensating wage equation is estimated as illustrated in (7.29). To compute the value of a statistical life, β_r is multiplied by 2,000 h per year to arrive at the annual compensation required for an additional unit of risk (assuming workers work 40 h per week for 50 weeks per year). The annual compensation is then multiplied by 10,000, the number of workers over which the risk is spread. In this example, an estimate of β_r equal to 0.35 would imply a VSL estimate of $7 million. Reviews of the early hedonic wage literature estimating the VSL are available in Mrozek and Taylor (2002), Viscusi and Aldy (2003), and Kochi et al. (2006).

In some environmental policies, particularly air quality policy, the benefits from mortality reductions can be the largest single component of the monetized benefits associated with the policy. For example, in the U.S. Environmental Protection Agency's (1997, 1999) retrospective (1970-1990) and prospective (1990-2010) analyses of the Clean Air Act, $4.8 million (1990 dollars) was the point estimate for the VSL used. This value was then aggregated over the number of lives saved to compute the monetized benefit of reduced mortality. In the retrospective analysis alone, the mean benefits of reduced mortality were estimated to be nearly $18 trillion (1990 dollars; U.S. Environmental Protection Agency 1997), just over 80% of the total benefits estimated to be associated with the Clean Air Act from 1970 to 1990.

From an empirical standpoint, many of the same issues considered for housing hedonic applications apply directly to labor markets. First, omitted variables and measurement error are as much a challenge in this literature, if not more so. Unobserved worker characteristics and job characteristics are likely correlated with job risks and wages and, therefore, bias the coefficient estimates for risk in an unknown direction (Black et al. 2003). While unobserved worker characteristics such as motivation or idiosyncratic job skills can be addressed within a panel-data framework (e.g., Kniesner et al. 2012), unobserved job characteristics are not.[28] Second, most studies rely on national average risk rates for broad occupational groups in broad industrial sectors. These measures are subject to considerable measurement error given their coarse resolution and the potential for dramatic differences among worksite safety practices even within the same industry (Lee and Taylor 2015).

Recently, quasi-experimental approaches have been used to address these concerns. Lee and Taylor (2015) provided the first quasi-experimental estimates

[28]Panel-data models that follow workers over time as they switch jobs (and thus, the risk of death on the job) do not control well for unobserved job characteristics because the models rely on job-changers to identify the marginal price of risk (i.e., β_r in Eq. (7.29)). To include individuals who do not change jobs in a panel-data framework, the researcher must rely on variation in workplace risks from year to year within the same job. Over relatively short periods of time, these variations are likely to be mostly random fluctuations in accidents. Over longer periods of time, changes in risk for any one specific job may be structural, but there is no reason to believe that there are not also changes in the correlated unobservables.

derived from a broad labor market setting. They employed federal workplace inspections that randomly select manufacturing plants for comprehensive safety inspections and use these inspections as an instrument for plant-level risks. The intuition behind the estimator is that prior to inspection, wages at riskier plants should be higher than those at less risky plants. Inspections provide an exogenous shock to the safety at plants that are inspected (an improvement), and wages are expected to fall (or rise less quickly) postinspection relative to uninspected plants whose safety remains unchanged. The estimator used to credibly identify wage-risk tradeoffs for manufacturing employees, and thus the VSL, is similar to the difference-in-differences estimator discussed in Sect. 7.2.3.2.[29]

A second important concern that has only partially been addressed within the hedonic framework is whether or not a VSL estimate derived from studies of fatal workplace risks is an appropriate value for reducing the risks that are regulated with environmental policy, such as the risk of premature death from exposure to poor air quality or cancer-causing agents (Cameron 2010; Scotton and Taylor 2011; Kochi and Taylor 2011). Deaths from cancers, illnesses, or heart attacks related to these environmental exposures may be associated with prolonged reductions in the quality of life prior to death or may involve more dread and pain as compared to workplace accidental deaths that are typically instantaneous. The author is unaware of research that estimates wage trade-offs workers make for on-the-job exposure to latent risks of death, such as exposures to cancer-causing agents.[30] This may be difficult to implement due to a lack of data and measurement difficulties but would be an important extension to the literature. In a related vein, a few studies have estimated the value of avoiding cancer-related illnesses within a housing hedonic context. Davis (2004) computed the implied estimate of a "VPL"—the value of (avoiding a case of) pediatric leukemia—to be \$5.6 million, and Gayer et al. (2000)

[29]Other quasi-experimental estimates of the VSL in nonlabor market settings also find lower VSL estimates. For example, Ashenfelter and Greenstone (2004) used changes in highway speed limits as a natural experiment. They first estimated the amount of saved travel time in states adopting increased speed limits and then estimated the impact of increased speed limits on rural interstate fatality rates. They reported an upper-bound point estimate for the VSL of \$2.2 million in 2011 dollars. León and Miguel (2013) also used a transportation setting to estimate the VSL, but within a developing country context. They used transportation mode choices (ferry, helicopter, hovercraft, and water taxi) and weather as an instrument for risk associated with each mode of travel to estimate a VSL of approximately \$600,000 for African travelers and approximately \$1 million for non-Africans.

[30]Scotton and Taylor (2011) and Kochi and Taylor (2011) attempted to address this issue by exploring the heterogeneity among workplace instantaneous deaths. Specifically, they note that nearly 20% of workplace deaths were due to homicide during the mid-1990s. Homicide risks have been shown to be viewed by individuals with more dread and fear than traditional job-related accidental risks, such as traffic and machinery accidents. Correspondingly, both studies find that homicide risks are compensated differently than traditional job-related accidental risks. While not conclusive, the authors argue that the results are suggestive that the VSLs derived from instantaneous workplace risks are not likely to be good estimates of the value of reducing risks associated with environmental exposures.

estimated the value of avoiding a cancer case to be \$4.3 million to \$5 million (all estimates are 2000 dollars).

A somewhat related concern—and one for which there is more research available—is whether or not the value of a statistical life derived from labor market studies based on working-age healthy adults is appropriate to extend to other populations at risk, such as the elderly or children. This is especially important for federal benefit-cost analysis of air quality regulations because much of the avoided mortality benefits of improved air quality accrues to the elderly population (for further discussion, see Eeckhoudt and Hammitt, 2001, Smith et al., 2004, Evans and Smith, 2006, and Aldy and Viscusi, 2007).

7.6 Summary

This chapter focused on the possibilities for nonmarket valuation using hedonic methods, primarily with property markets. Estimation of the hedonic price function alone (a first-stage analysis) is by far the most common hedonic application. Indeed, a search using a popular academic search engine using the keyword "hedonic" revealed more than 850 published works in the economics literature. The popularity of the hedonic method is not surprising given the relatively minimal data requirements and straightforward empirical implementation. In addition, the insights that may be drawn from a first-stage analysis are quite powerful: the value of a change in an environmental amenity that accrues to a population of residents in an area. But, it was also demonstrated that many details require careful consideration when estimating the hedonic price function. Current hedonic applications are expected to have intensive data and estimation procedures. Researchers must carefully consider their data sources, choices of the independent variables and their measurement, the functional form of the price function, the sample frame used, and the estimation procedures employed. While not all issues will require in-depth treatment, the case must be made that the researcher has identified without bias the implicit price they seek to measure.

Even though very many first-stage hedonic analyses have been conducted, this is still an exciting area of research. Advances in computing and explosive growth in freely available geospatial data have resulted in spatial modeling being fully integrated into empirical analyses. Quasi-experimental designs for evaluating the impact of changing environmental amenities on property prices have also grown rapidly in recent years, and this too will be an area for expanding research.

Beyond the first-stage estimation of the hedonic price function, much additional research can be conducted on estimating welfare changes within the hedonic framework. This is an important component of the hedonic nonmarket valuation literature because many policy changes are not localized in context, thus requiring demand analyses. Although welfare analysis in this context has been actively discussed since Rosen's (1974) seminal article, it has remained relatively rare in implementation. More recently, however, sorting models have gained traction as an alternative to the data-intensive multiple-market approach to estimating demand. Here again, as rich data on properties *and* their owners are ever increasing in their availability, the scope for demand analyses can only increase.

Acknowledgements The author would like to thank Kevin Boyle, Tom Brown, Patty Champ, Nicolai Kuminoff, Jeffrey Zabel, and the participants in the nonmarket valuation workshop sponsored by the Rocky Mountain Research Station, April 2013, for helpful comments.

References

Aldy, J. E. & Viscusi, W. K. (2007). Age differences in the value of statistical life: Revealed preference evidence. Review of Environmental Economics and Policy, 1, 241-260.

Anselin, L. (1988). Spatial econometrics: Methods and models. Dordrecht, The Netherlands: Kluwer Academic.

Ashenfelter, O. & Greenstone, M. (2004). Using mandated speed limits to measure the value of a statistical life. Journal of Political Economy, 112, S226-S267.

Bajari, P., Fruehwirth, J. C., Kim, K. I. & Timmins, C. (2012). A rational expectations approach to hedonic price regressions with time-varying unobserved product attributes: The price of pollution. American Economic Review, 102, 1898-1926.

Bajari, P. & Kahn, M. E. (2005). Estimating housing demand with an application to explaining racial segregation in cities. Journal of Business & Economic Statistics, 23, 20-33.

Baranzini, A. & Ramirez, J. V. (2005). Paying for quietness: The impact of noise on Geneva rents. Urban Studies, 42, 633-646.

Baranzini, A. & Schaerer, C. (2011). A sight for sore eyes: Assessing the value of view and land use in the housing market. Journal of Housing Economics, 20, 191-199.

Bartik, T. J. (1986). Neighborhood revitalization's effects on tenants and the benefit-cost analysis of government neighborhood programs. Journal of Urban Economics, 19, 234-248.

Bartik, T. J. (1987). The estimation of demand parameters in hedonic price models. Journal of Political Economy 95:81-88.

Bartik, T. J. (1988). Measuring the benefits of amenity improvements in hedonic price models. Land Economics, 64, 172-183.

Bin, O., Crawford, T. W., Kruse, J. B. & Landry, C. E. (2008). Viewscapes and flood hazard: Coastal housing market response to amenities and risk. Land Economics, 84, 434-448.

Bin, O. & Landry, C. E. (2013). Changes in implicit flood risk premiums: Empirical evidence from the housing market. Journal of Environmental Economics and Management, 65, 361-376.

Black, D., Galdo, J. & Liu, L. (2003). How robust are hedonic wage estimates of the price of risk: The final report. Report submitted to U.S. Environmental Protection Agency, National Center for Environmental Economics, Washington, DC. http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0483-04.pdf/$file/EE-0483-04.pdf.

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. The Quarterly Journal of Economics, 114, 577-599.

Bockstael, N. E. & McConnell, K. E. (2007). Environmental and resource valuation with revealed preferences: A theoretical guide to empirical models. Vol. 7 of I. Bateman (Ed.), The economics of non-market goods and resources. Dordrecht, The Netherlands: Springer.

Boes, S. & Nüesch, S. (2011). Quasi-experimental evidence on the effect of aircraft noise on apartment rents. Journal of Urban Economics, 69, 196-204.

Bound, J., Jaeger, D. A. & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. Journal of the American Statistical Association, 90, 443-450.

Boyle, K., Lewis, L., Pope, J. & Zabel, J. (2012). Valuation in a bubble: Hedonic modeling pre- and post-housing market collapse. AERE Newsletter, 32 (2), 24-32.

Boyle, K. J., Poor, P. J. & Taylor, L. O. (1999). Estimating the demand for protecting freshwater lakes from eutrophication. American Journal of Agricultural Economics, 81, 1118-1122.

Brady, M. & Irwin, E. (2011). Accounting for spatial effects in economic models of land use: Recent developments and challenges ahead. Environmental and Resource Economics, 48, 487-509.

Brander, L. M. & Koetse, M. J. (2011). The value of urban open space: Meta-analyses of contingent valuation and hedonic pricing results. Journal of Environmental Management, 92, 2763-2773.

Buck, S., Aufhammer, M. & Sunding, D. (2014). The Value of Heterogeneous Property Rights and the Costs of Water Volatility. American Journal of Agricultural Economics, 96, 953–969.

Cameron, T. A. (2010). Euthanizing the value of a statistical life. Review of Environmental Economics and Policy, 4, 161-178.

Campbell, J. Y., Giglio, S. & Pathak, P. (2011). Forced sales and house prices. American Economic Review, 101, 2108-2131.

Carson, R. T. & Dastrup, S. R. (2013). After the fall: An ex post characterization of housing price declines across metropolitan areas. Contemporary Economic Policy, 31, 22-43.

Chattopadhyay, S. (1999). Estimating the demand for air quality: New evidence based on the Chicago housing market. Land Economics, 75, 22-38.

Chay, K. Y. & Greenstone, M. (2005). Does air quality matter? Evidence from the housing market. [Electronic version]. Journal of Political Economy, 113, 376-424.

Cheshire, P. & Sheppard, S. (1995). On the price of land and the value of amenities. Economica, 62, 247-267.

Cheshire, P. & Sheppard, S. (1998). Estimating the demand for housing, land, and neighbourhood characteristics. Oxford Bulletin of Economics and Statistics, 60, 357-382.

Cohen, J., Blinn, C. E., Boyle, K. J., Holmes, T. P. & Moeltner, K. (2014). Hedonic valuation with translating ameniities: Mountain pine beetles and host trees in the Colorado front range. Environmental and Resource Economics, 59, December.

Coulson, N. E & Zabel, J. E. (2013). What can we learn from hedonic models when housing markets are dominated by foreclosures? Annual Review of Resource Economics, 5, 261-279.

Cropper, M. L., Deck, L. B., Kishor, G. & McConnell, K. E. (1993). The estimation of consumer preferences for attributes: A comparison of hedonic and discrete choice approaches. The Review of Economics and Statistics, 75, 225-232.

Cropper, M. L., Deck, L. B. & McConnell, K. E. (1988). On the choice of functional form for hedonic price functions. The Review of Economics and Statistics 70, 668-675.

Davis, L. W. (2004). The effect of health risk on housing values: Evidence from a cancer cluster. American Economic Review, 94, 1693-1704.

Davis, L.W. (2011). The effect of power plants on local housing values and rents. Review of Economics and Statistics, 93, 1391-1402.

Eeckhoudt, L. R. & Hammitt J. K. (2001). Background risks and the value of a statistical life. Journal of Risk and Uncertainty, 23, 261-279.

Epple, D., Quintero, L. & Sieg, H. (2013). Estimating hedonic functions for rents and values in the presence of unobserved heterogeneity in the quality for housing. GSIA Working Papers Series No. 2013-E34. Electronic version. Pittsburgh, PA: Carnegie Mellon Tepper School of Business.

Evans, M. F. & Smith, V. K. (2006). Do we really understand the age-VSL relationship? Resource and Energy Economics, 28, 242-261.

Freeman, A. M., III, Herriges, J. A. & Kling, C. L. (2014). The measurement of environmental and resource values: Theory and Methods. Washington, DC: RFF Press.

Gamper-Rabindran, S. & Timmins, C. (2013). Does cleanup of hazardous waste sites raise housing values? Evidence of spatially localized benefits. Journal of Environmental Economics and Management, 65. 345-360.

Garber, S. & Klepper, S. (1980). Extending the classical normal errors-in-variables model. Econometrica, 48, 1541-1546.

Gawande, K., Jenkins-Smith, H. & Yuan, M. (2013). The long-run impact of nuclear waste shipments on the property market: Evidence from a quasi-experiment. Journal of Environmental Economics and Management, 65, 56-73.

Gayer, T., Hamilton, J. T. & Viscusi, W. K. (2000). Private values of risk tradeoffs at superfund sites: Housing market evidence on learning about risk. Review of Economics and Statistics, 82, 439-451.

Goodman, A. C. & Thibodeau, T. G. (2007). The spatial proximity of metropolitan area housing submarkets. Real Estate Economics, 35, 209-232.

Gopalakrishnan, S. & Klaiber, H. A. (2014). Is the shale energy boom a bust for nearby residents? Evidence from housing values in Pennsylvania. American Journal of Agricultural Economics, 96, 43-66. DOI:10.1093/ajae/aat065.

Grainger, C. A. (2012). The distributional effects of pollution regulations: Do renters fully pay for cleaner air? Journal of Public Economics, 96, 840-852.

Greene, W. H. (2011). Econometric Analysis (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Greenstone, M. & Gallagher, J. (2008). Does hazardous waste matter? Evidence from the housing market and the superfund program. The Quarterly Journal of Economics, 123, 951-1003.

Hallstrom, D. G. & Smith, V. K. (2005). Market responses to hurricanes. Journal of Environmental Economics and Management, 50, 541-561.

Halvorsen, R. & Palmquist, R. (1980). The interpretation of dummy variables in semilogarithmic equations. American Economic Review, 70, 474-475.

Hanna, B. G. (2007). House values, incomes, and industrial pollution. Journal of Environmental Economics and Management, 54, 100-112.

Hausman, J. A. (1978). Specification tests in econometrics. Econometrica, 46, 1251-1271.

Hausman, J. A. (1981). Exact consumer's surplus and deadweight loss. American Economic Review, 71, 662-676.

Heintzelman, M. D. & Tuttle, C. M. (2012). Values in the wind: A hedonic analysis of wind power facilities. Land Economics, 88, 571-588.

Hess, D. B. & Almeida, T. M. (2007). Impact of proximity to light rail rapid transit on station-area property values in Buffalo, New York. Urban Studies, 44, 1041-1068.

Hite, D., Chern, W., Hitzhusen, F. & Randall, A. (2001). Property-value impacts of an environmental disamenity: The case of landfills. Journal of Real Estate Finance and Economics, 22, 185-202.

Huang, J.-C. & Palmquist, R. B. (2001). Environmental conditions, reservation prices, and time on the market for housing. Journal of Real Estate Finance and Economics, 22, 203-219.

Ihlanfeldt, K. R. & Taylor, L. O. (2004). Externality effects of small-scale hazardous waste sites: Evidence from urban commercial property markets. Journal of Environmental Economics and Management, 47, 117-139.

Irwin, E. G. & Bockstael, N. E. (2001). The problem of identifying land use spillovers: Measuring the effects of open space on residential property values. American Journal of Agricultural Economics, 83, 698-704.

Kennedy, P. E. (1981). Estimation with correctly interpreted dummy variables in semilogarithmic equations. American Economic Review, 72, 801.

Kiel, K. A. & McClain, K. T. (1995). Housing prices during siting decision stages: The case of an incinerator from rumor through operation. Journal of Environmental Economics and Management, 28, 241-255.

Kiel, K. A. & Williams, M. (2007). The impact of Superfund sites on local property values: Are all sites the same? Journal of Urban Economics, 61, 170-192.

Kiel, K. A. & Zabel, J. E. (1999). The accuracy of owner-provided house values: The 1978-1991 American Housing Survey. Real Estate Economics, 27, 263-398.

Kim, C. W., Phipps, T. T. & Anselin, L. (2003). Measuring the benefits of air quality improvement: A spatial hedonic approach. Journal of Environmental Economics and Management, 45, 24-39.

Kim, J. & Goldsmith, P. (2009). A spatial hedonic approach to assess the impact of swine production on residential property values. Environmental and Resource Economics, 42, 509-534.

Klaiber, H. A. & Kuminoff, N. V. (2014). Equilibrium sorting models of land use and residential choice. In J. M. Duke & J. Wu (Eds.), The Oxford handbook of land economics (pp. 352-379). Oxford, United Kingdom: Oxford University Press.

Klaiber, H. A. & Phaneuf, D. J. (2010). Valuing open space in a residential sorting model of the Twin Cities. Journal of Environmental Economics and Management, 60, 57-77.

Klaiber, H. A. & Smith, V. K. (2013). Quasi experiments, hedonic models, and estimating trade-offs for local amenities. Land Economics, 89, 413-431.

Kniesner, T. J., Viscusi, W. K., Woock, C. & Ziliak, J. P. (2012). The value of a statistical life: Evidence from panel data. The Review of Economics and Statistics, 94, 74-87.

Kochi I., Hubbell, B. & Kramer, R. (2006). An empirical Bayes approach to combining and comparing estimates of the value of a statistical life for environmental policy analysis. Environmental & Resource Economics, 34, 385-406.

Kochi, I. & Taylor, L. O. (2011). Risk heterogeneity and the value of reducing fatal risks: Further market-based evidence. Journal of Benefit-Cost Analysis, 2, 1-28.

Kohlhase, J. E. (1991). The impact of toxic waste sites on housing values. Journal of Urban Economics, 30, 1-26.

Kuminoff, N. V., Parmeter, C. F. & Pope, J. C. (2010). Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? Journal of Environmental Economics and Management, 60, 145-160.

Kuminoff, N. V. & Pope, J. C. (2012). A novel approach to identifying hedonic demand parameters. Economics Letters, 116, 374-376.

Kuminoff, N. V. & Pope, J. C. (2013). The value of residential land and structures during the great housing boom and bust. Land Economics, 89, 1-29.

Kuminoff, N. V. & Pope, J. C. (2014). Do 'capitalization effects' for public goods reveal the public's willingness to pay? International Economic Review, 55, 1227-1250.

Kuminoff, N. V., Smith, V. K. & Timmins, C. (2013). The new economics of equilibrium sorting and policy evaluation using housing markets. Journal of Economic Literature, 51, 1007-1062.

Landry, C. E. & Hindsley, P. (2011). Valuing beach quality with hedonic property models. Land Economics, 87, 92-108.

Leamer, E. (2007). Housing IS the business cycle. Proceedings, Federal Reserve Bank of Kansas City (pp. 149-233).

Lee, J. & Taylor, L. O. (2015). Randomized safety inspections and risk exposure on the job: quasi-experimental estimates of the value of a statistical life. Working Paper. Center for Environmental and Resource Economic Policy, North Carolina State University.

Leggett, C. G. & Bockstael, N. E. (2000). Evidence of the effects of water quality on residential land prices. Journal of Environmental Economics and Management, 39, 121-144.

León, G. & Miguel, E. (2013). Transportation choices and the value of statistical life. National Bureau of Economic Research Working Paper 19494. Cambridge, MA: NBER. Last accessed Nov. 21, 2013, from http://emiguel.econ.berkeley.edu/research/transportation-choices-and-the-value-of-statistical-life-in-africa.

LeSage, J. & Pace, R. K. (2009). Introduction to spatial econometrics. Boca Raton, FL: Chapman & Hall/CRC.

Linneman, P. (1980). Some empirical results on the nature of the hedonic price function for the urban housing market. Journal of Urban Economics 8, 47-68.

Liu, X. & Lynch, L. (2011). Do agricultural land preservation programs reduce farmland loss? Evidence from a propensity score matching estimator. Land Economics, 87, 183-201.

Liu, X., Taylor, L. O., Hamilton, T. L. & Grigelis, P. E. (2013). Amenity values of proximity to National Wildlife Refuges: An analysis of urban residential property values. Ecological Economics, 94, 37-43.

Matthews, J. W. & Turnbull, G. K. (2007). Neighborhood street layout and property value: The interaction of accessibility and land use mix. Journal of Real Estate Finance and Economics, 35, 111-141.

McConnell, V. & Walls, M. (2005). The value of open space: Evidence from studies of nonmarket benefits. Washington, DC: Resources for the Future.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. Journal of Business and Economic Statistics, 13, 151-61.

Moffit, R. (1991). Program evaluation with nonexperimental data. Evaluation Review, 15, 291-314.

Mrozek, J. R. & Taylor, L. O. (2002). What determines the value of life? A meta-analysis. Journal of Policy Analysis and Management, 21, 253-270.

Nakamura, A. & Nakamura, M. (1998). Model specification and endogeneity. Journal of Econometrics, 83, 213-237.

Noonan, D. S., Krupka, D. J. & Baden, B. M. (2007). Neighborhood dynamics and price effects of superfund site clean-up. Journal of Regional Science, 47, 665-692.

Ohta, M. & Griliches, Z. (1976). Automobile prices revisited: Extensions of the hedonic hypothesis. In N. Terleckyj (Ed.), Household production and consumption (pp. 325-398). Cambridge, MA: National Bureau of Economic Research.

Palmquist, R. B. (1984). Estimating the demand for the characteristics of housing. The Review of Economics and Statistics, 66, 394-404.

Palmquist, R. B. (1988). Welfare measurement for environmental improvements using the hedonic model: The case of nonparametric marginal prices. Journal of Environmental Economics and Management, 15, 297-312.

Palmquist, R. B. (1991). Hedonic methods. In J. B. Braden & C. D. Kolstad (Eds.), Measuring the demand for environmental quality (pp. 77-120). North-Holland, Amsterdam, The Netherlands: North-Holland.

Palmquist, R. B. (1992a). Valuing localized externalities. Journal of Urban Economics, 31, 59-68.

Palmquist, R. B. (1992b). A note on transactions costs, moving costs, and benefit measurement. Journal of Urban Economics, 32, 40-44.

Palmquist, R. B. (2006). Property value models. In K.-G. Mäler & J. R. Vincent (Eds.), Handbook of environmental economics, Volume 2: Valuing environmental changes (pp. 763-819). Amsterdam, The Netherlands: Elsevier.

Palmquist, R. B. & Israngkura, A. (1999). Valuing air quality with hedonic and discrete choice models. American Journal of Agricultural Economics, 81, 1128-1133.

Parmeter, C. F. & Pope, J. C. (2013). Quasi-experiments and hedonic property value methods. In J. A. List & M. K. Price (Eds.), Handbook of experimental economics and the environment (pp. 3-65). Cheltenham, United Kingdom: Edward Elgar.

Parsons, G. R. (1986). An almost ideal demand system for housing attributes. Southern Economic Journal, 53, 347-363.

Paterson, R. W. & Boyle, K. J. (2002). Out of sight, out of mind? Using GIS to incorporate visibility in hedonic property value models. Land Economics, 78, 417-425.

Petrie, R. A. & Taylor, L. O. (2007). Estimating the value of water use permits: A hedonic approach applied to farmland in the southeastern United States. Land Economics, 83, 302-318.

Phaneuf, D.J. & Requate, T. (in press). A course in enviornmental economics: Theory, policy, and practice. New York: Cambridge University Press.

Phaneuf, D. J., Taylor, L. O. & Braden, J. B. (2013). Combining revealed and stated preference data to estimate preferences for residential amenities: A GMM approach. Land Economics, 89, 30-52.

Poor, P. J., Boyle, K. J., Taylor, L. O. & Bouchard, R. (2001). Objective versus subjective measures of water quality in hedonic property value models. Land Economics, 77, 482-493.

Pope, J. C. (2008). Buyer information and the hedonic: The impact of a seller disclosure on the implicit price for airport noise. Journal of Urban Economics, 63, 498-516.

Quigley, J. M. (1982). Nonlinear budget constraints and consumer demand: An application to public programs for residential housing. Journal of Urban Economics, 12, 177-201.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. Journal of Political Economy 82, 34-55.

Sander, H., Polasky, S. & Haight, R. G. (2010). The value of urban tree cover: A hedonic property price model in Ramsey and Dakota Counties, Minnesota, USA. Ecological Economics, 69, 1646-1656.

Sanders, N. J. (2012). Toxic assets: How the housing market responds to environmental information shocks. Working Paper 128. Williamsburg, VA: College of William and Mary, Department of Economics. Retrieved Nov. 2, 2103, from IDEAS database.

Scotton, C. R. & Taylor, L. O. (2011). Valuing risk reductions: Incorporating risk heterogeneity into a revealed preference framework. Resource and Energy Economics, 33, 381-397.

Shiller, R. J. (2007). Understanding recent trends in house prices and home ownership. National Bureau of Economic Research Working Paper 13553. Cambridge, MA: NBER.

Sieg, H., Smith, V. K., Banzhaf, H. S. & Walsh, R. (2004). Estimating the general equilibrium benefits of large changes in spatially delineated public goods. International Economic Review, 45, 1047-1077.

Small, K. A. & Steimetz, S. S. C. (2012). Spatial hedonics and the willingness to pay for residential amenities. Journal of Regional Science, 52, 635-647.

Smith, V. K., Evans, M. F., Kim, H. & Taylor, D. H. Jr. (2004). Do the near-elderly value mortality risks differently? Review of Economics and Statistics, 86, 423-29.

Smith, V. K. & Huang, J.-C. (1995). Can markets value air quality? A meta-analysis of hedonic property value models. Journal of Political Economy, 103, 209-227.

Taylor, L. O. (2008). Theoretical foundations and empirical developments in hedonic modeling. In A. Baranzini, J. Ramirez, C. Schaerer & P. Thalmann (Eds.), Hedonic methods in housing markets: Pricing environmental amenities and segregation (pp. 15-37). New York: Springer.

Taylor, L. O. Phaneuf, D. J. & Liu, X. (2016). Disentangling the property value impacts of environmental contamination from locally undesirable land uses: Implications for measuring post-cleanup stigma, Journal of Urban Economics, 93, 85-98.

Taylor, L. O. & Smith, V. K. (2000). Environmental amenities as a source of market power. Land Economics, 76, 550-568.

Todd, P. E. (2008). Matching estimators. From: The New Palgrave Dictionary of Economics. In S. N. Durlauf & L. E. Blume (Eds.). The New Palgrave Dictionary of Economics Online (2nd ed.). Palgrave Macmillan. Retrieved from www.dictionaryofeconomics.com/article?id=pde2008_M000365.

Tra, C. I. (2010). A discrete choice equilibrium approach to valuing large environmental changes. Journal of Public Economics, 94, 183-196.

Tra, C. I. (2013). Measuring the general equilibrium benefits of air quality regulation in small urban areas. Land Economics, 89, 291-307.

U.S. Environmental Protection Agency. (1997). The benefits and costs of the Clean Air Act: 1970 to 1990. Prepared by the Office of Administration and Resources Management, Office of Policy, Planning, and Evaluation for the U.S. Congress, October 1997.

U.S. Environmental Protection Agency. (1999). The benefits and costs of the Clean Air Act: 1990 to 2010. Prepared by the Office of Air and Radiation, Office of Policy for the U.S. Congress, November 1999.

Vartia, Y. O. (1983). Efficient methods of measuring welfare change and compensated income in terms of ordinary demand functions. Econometrica, 51, 79-98.

Viscusi, W. K. (1992). Fatal tradeoffs: Public and private responsibilities for risk. New York: Oxford University Press.

Viscusi, W. K. (1993). The value of risks to life and health. Journal of Economic Literature, 31, 1912-1946.

Viscusi, W. K. & Aldy, J. E. (2003). The value of a statistical life: A critical review of market estimates throughout the world. Journal of Risk and Uncertainty, 27, 5-76.

Waugh, F. V. (1928). Quality factors influencing vegetable prices. Journal of Farm Economics, 10, 185-196.

Zabel, J. (2013). Hedonic models and the housing cycle. Working Paper, Department of Economics. Medford, MA: Tufts University.

Zabel, J. E. & Guignet, D. (2012). A hedonic analysis of the impact of LUST sites on house prices. Resource and Energy Economics, 34, 549-564.

Zabel, J. E., and Kiel, K. A. (2000). Estimating the demand for air quality in four U.S. cities. Land Economics, 76, 174-194.